ARBITRAGE IN STOCK INDEX FUTURES ONE AND TWO DIMENSIONAL PROBLEMS

WANG SHENGYUAN

(B.Sci.(Hons.). NUS)

A THESIS SUBMITTED

FOR THE DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF MATHEMATICS

NATIONAL UNIVERSITY OF SINGAPORE

2009

Acknowledgement

First of all, sincere gratitude is extended to my supervisor, Professor Dai Min. I have benefited greatly from his considerable help and guidance. I will always remember for his insightful supervision and earnest backing all through the searching, analysis and paper-writing stages. It would be impossible for me to reach the level of this paper without his instruction. Most importantly, it is beneficial to my whole life.

Of course, people too numerous to mention who have made my undergraduate study at NUS both productive and enjoyable. I also want to thank Professor Yue-Kuen Kwok, who has been pleased to share his expertise on this topic. Various friends have helped me to conquer problems, both real and imagined. Particular mention must be made of Li Pei Fan. She is a such kind senior who always gives me many valuable issues on numerical methods and helps me check on Matlab code. I would be also thankful for Zhong Yi Fei for validating my numerical results and pointing out my mistakes. Members of the math lab, both past and present, have always been there when needed. And a heartfelt thanks to all who have helped me in one way or another.

Last, and certainly not least, I am extremely thankful to my girlfriend and parents, for their support and patience over these two year. Specially to my girlfriend, may we grow closer together as I finally move past the student phase of life. Also thanks to my colleagues and bosses at Octagon Advisors: CEO Koh Beng Seng, MD David Loh, Director Chiah Kok Hoe, Director Chan Chin Hiang and Director Chng Say Keong for their undersanding of my constraints, because of, academic reasons.

Contents

Summary

Stock indexes, unlike stocks, options, cannot be trader directly, so futures based on stock indexes are primary way of trading stock indexes. There are three type of investors in various financial markets, namely, speculator, hedger and arbitrager. In this thesis, we are interested in the arbitrage profit in stock index futures. This thesis mainly focus on on pricing options whose payoff is based on simple arbitrage profit in stock index futures and plotting their early exercise boundaries. We consider both one dimensional and two dimensional problems, for each we sub-divide as 'no position limits' case and 'with position limits' case.

In one dimensional problem, we use Brownian Bridge process to model simple arbitrage profit. A one dimensional PDE for the options is derived. In two dimensional problem, we add one mean-reverting stochastic differential equation to model order imbalance. A two dimensional PDE for the options is derived. We also take into account of transaction costs and position limits and form complete models.

For numerical experiement, we use fully implicit and Crank-Nicolson scheme to solve the variational inequality numerically. To handle American option type, we adopt projected SOR method. Numerical Results of the early exercise boundaries and option values are given and analyzed. These early exercise boundaries give us the optimal arbitrage strategy. We discuss various parameter effects on option values and early exercise boundary, for one dimensional problem, while we also examine the order imbalance impacts on early exercise boundary, for two dimensional problem. We also compare the numerical results between the 'no position limits' and 'with position limits' models, and find the optimal trading strategy is exactly the same for both cases.

Keywords: stock index futures, simple arbitrage profit, order imbalance, optimal trading strategy.

List of Tables

List of Figures

List of Algorithms

Chapter 1

Introduction

1.1 Background

1.1.1 Arbitrage in Stock Index Futures

The textbook definition of arbitrage suggests that it is a straightforward matter of taking offsetting positions in different securities and realizing the riskless profit. It can be achieved by either taking advantage of price discrepancies of the same product in different financial market, or by deriving more complicated strategies to earn the arbitrage profit - such as in the stock index futures case.

Index futures are futures markets where the underlying commodity is a stock index, such as the DJIA, S&P, or the FTSE100[1]. Stock indexes cannot be traded directly, so futures based upon stock indexes are primary way of trading stock indexes. Index futures are essentially the same as all other futures markets, like currency and commodity futures markets, and are traded in exactly the same way.

A stock index futures is a forward contract to obtain a stock index on the settlement date of the contract. To derive a general theoretical arbitrage relation between spot

[1]DJIA: Dow Jones Industrial Average. S&P: Standard and Poor. FSTE: Financial Times Stock Exchange

and futures prices, consider a futures contract of maturity T. Let $F_t(T)$ be the futures price at maturity date, $P_t(T)$ be the price of a $T-t$ period unit discount bond, and S_t be the current spot price of the underlying portfolio. Define

$$G_t := F_t(T) \cdot P_t(T) + \mathrm{PV}(\mathrm{div})$$

where PV(div) is the present value of the dividends payable on the underlying portfolio up to the maturity of the contract. Denote ϵ as the arbitrage profit in the absence of transaction costs to be obtained by taking a long position in the underlying portfolio, hedging it with a short position in the futures contract, and holding the position until maturity of the futures contract: we shall refer to this as a *simple* long arbitrage position; it is simple because it is to be held until maturity. Then

$$\epsilon = G_t - S_t$$

The strategy is to borrow an amount of G_t and to buy one unit of the underlying portfolio at cost S_t. By constructing G_t in this pattern, this strategy yields an immediate cash inflow of ϵ and no further net cash flows. To confirm this point, let us check what will happen at maturity date. We need pay off the loan that we have borrowed at initial time. The amount we need to pay is

$$G_T = \frac{G_t}{P_t(T)} = \frac{F_t(T) \cdot P_t(T) + \mathrm{PV}(\mathrm{div})}{P_t(T)} = F_t(T) + \mathrm{FV}(\mathrm{div})$$

However, at maturity date, we exercise the futures contract to sell the underlying portfolio at future price $F_t(T)$ which will pay off part of the loan, the balance FV(div) being paid is received for holding the underlying portfolio. Essentially, there is no cash flow involved after initial time. Therefore, ϵ is the value of the arbitrage profit to be reaped from this simple long arbitrage position.

Note that, if ϵ is negative, we can reverse the above strategy to obtain an arbitrage profit of $-\epsilon$. The strategy is to deposit an amount of G_t and to short one unit of the underlying portfolio as cost S_t. Similarly, we will gain some amount of money at maturity date,

$$G_T = \frac{G_t}{P_t(T)} = \frac{F_t(T) \cdot P_t(T) + \text{PV(div)}}{P_t(T)} = F_t(T) + \text{FV(div)}$$

However, we use part of the gain, FV(div), to pay for shorting the underlying portfolio, and we need to exercise the futures contract, buying back the underlying portfolio at future price $F_t(T)$ to close the short position. Therefore, $-\epsilon$ is the value of the arbitrage profit to be reaped from this simple short arbitrage position.

1.1.2 Transaction Costs

Since stock index arbitrage involves transactions in both the stock and futures markets, account must be taken of commissions and bid-ask spreads in the two markets. To open an arbitrage position involves a future commission, a stock commission, and the market impact associated with the stock transaction, due to the bid-ask spread. If the arbitrage position is held to expiration, the only additional cost is the commission to close out the futures position and the stock commission associated with the reversal of the stock position. No market-impact costs are incurred since the stock can be sold at the market closing price, which is the same as the terminal futures price. However, if the arbitrage position is closed out early, there is an additional cost consisting of the market-impact cost on the stock position.

Therefore, we use C_1 and C_2 to denote the costs associated with the simple arbitrage and the incremental costs associated with early close out, namely

$$\begin{cases} C_1 = \text{two futures commissions} + \text{two stock commissions} + \text{one market-impact cost} \\ C_2 = \text{one market-impact cost} \end{cases}$$

1.2 Historical Work And Author's Contribution

Numerous famous academicians and practitioners have done extensive research on stock index futures. We present the major historical works in a chronological order.
In [1], Bradford Cornell and Kenneth R.French suggest the discrepancy between the actual and predicted stock index futures prices is caused by taxes. The fact that capital gains and losses are not taxed until they are realized gives stockholders a valuable timing option. Since this option is not available to stock index futures traders, the futures prices will be lower than standard no-tax models predict.
In [2], Figlewski finds that the standard deviation of daily returns on portfolio regarding to NYSE[2] Index, hedged by a short position in the nearest NYSE futures contract, was 19.72% during January 1981 to March 1982. The corresponding figure for S&P 500 portfolio for the same period was 16.46%. These numbers show these contracts do not always trade at the prices predicted by a simple arbitrage relation with the spot price.
In [3], Michael J. Brennan and Eduardo S. Schwartz uses a continuous-time Brownian Bridge to model the stochastic process of simple arbitrage profit, and proposes a PDE approach for pricing the options whose underlying is the simple arbitrage profit.
In [4], Joseph K.W. Fung introduces order imbalance as measure of both the direction and the extent of market liquidity. The study covers the period of the Asian financial crisis and includes wide variations in order imbalance and the index futures basis. The results indicate that the arbitrage spread is positively related to the aggregate order imbalance in the underlying index stocks, and negative order imbalance has stronger impact than positive order imbalance.
In [5], Joseph K.W. Fung and Philip L.H Yu uses transaction records of index futures and index stocks, with bid/ask price quotes, to examine the impact of stock

[2]NYSE: New York Stock Exchange

market order imbalance on the dynamic behavior of index futures and cash index prices. Their findings indicate that a stock market microstructure that allows a quick resolution of order imbalance promotes dynamic arbitrage efficiency between futures and underlying stocks.

In [6], Chen Huan uses explicit method to price one dimensional options and draw their respective early exercise boundaries. Convergence of the model is also analyzed. In [7], Dai Kwok and Zhong use one mean-reverting stochastic differential equation to model order imbalance and give me the motivation to price options by a two dimensional PDE.

The main contributions of this thesis are

- We carry out a two dimensional PDE approach to solve the option values numerically. We adopt a fully implicit and Crank Nicolson scheme, where central differencing is used as much as possible. Upwinding scheme is also used to ensure the row diagonal dominance of M-matrix. We handle the American option type with projected SOR method.
- We discuss various parameter effects on option values and early exercise boundary, for one dimensional problem, while we also examine the order imbalance impacts on early exercise boundary, for two dimensional problem.
- We compare the numerical results between the 'no position limits' and 'with position limits' models, and find the optimal trading strategy is exactly the same for both cases.

1.3 Outline

The thesis is mainly motivated by the paper [3] and [4]. In [3], a PDE approach is adopted to price the options whose underlying is simple arbitrage profit. It is a one dimensional problem. In [4], the concept of order imbalance, which clearly has an impact on the options price, is introduced. In this thesis, beyond the historical works, we are going to build the option model on simple arbitrage profit and order imbalance[3], derive its govern PDE, evaluate the option price and plot the early exercise regions or boundaries by numerical methods.

Chapter 1 gives you some fundamental understanding on the arbitrage in stock index futures market. The remainder of this thesis is organized as follows. In chapter 2, we derive the PDE for option on one simple arbitrage profit, use project SOR with fully implicit and Crank-Nicolson method to evaluate option prices numerically, and also present the plot of early exercise regions and boundaries. Additionally, we discuss the parameter effects on options price and early exercise boundaries. In chapter 3, we introduce order imbalance in the stock futures market, and extend to two dimensional case, namely, the value of option depending on simple arbitrage profit and order imbalance. The numerical algorithms are provided and the plot of option values and early exercise boundary are presented. In chapter 4, we design options on two simple arbitrage profit with various payoff types. Finally, concluding remarks and possible future research direction are drawn in chapter 5. The Matlab source code is not given in Appendix due to the large size, and is packaged as an external file.

[3]It is a two dimensional problem

Chapter 2

One Dimensional Problem

2.1 Theoretical Model

In this section we focus on one dimensional problem and derive the partial differential equation for the options to close out or initiate a stock index arbitrage position, and construct the complete model for 'no position limits' case and 'with position limits' case.

2.1.1 Underlying Asset and Options

A simple long arbitrage position as defined involves a long position in the underlying portfolio and a short position in the futures contract, held to maturity. ϵ is the riskless profit obtained by establishing such a position. Similarly, we define a simple short arbitrage position as a short position in the underlying portfolio and a long position in the futures contract, held to maturity. $-\epsilon$ is the riskless profit from establishing such a position.

Technically speaking, a long (short) arbitrage position can be closed-out prior to maturity by taking an offsetting short (long) arbitrage position. Without regarding to transaction costs, this immediately yields an additional arbitrage profit of

$-\epsilon(\epsilon)$.

Let $V(\epsilon,t)(U(\epsilon,t))$ be the value of the right to close a long (short) arbitrage position prior to maturity when the simple arbitrage profit before transaction costs is ϵ and the time to maturity of the futures contract is $T-t$. Similarly, let $W(\epsilon,t)$ be the value of the right to initiate an arbitrage position.

In order to value the arbitrage and early close-out options and determine the optimal strategies for exercising them, it is necessary to make some assumptions about the stochastic differential equation (SDE) of ϵ. We assume that the simple arbitrage profit follows a continuous-time Brownian Bridge process.

$$d\epsilon = -\frac{\mu\epsilon}{T-t}dt + \gamma dW \tag{2.1}$$

Some explanations on these parameters

$T-t$ is the time to maturity of the futures contract

μ is the speed of mean reversion

γ is the instantaneous standard deviation of the process

dW is the increment to a Gauss-Wiener process

The Brownian Bridge process has the property that the arbitrage profit tends to be zero and is zero at maturity almost surely. It makes economical sense because when close to maturity, the mean-reverting parameter $\frac{\mu}{T-t}$ is quite large, ϵ will act so quickly as to bring the variable back to its mean level, namely zero, as arbitragers will always take existing arbitrage opportunities to drive the profit to zero[1].

By risk neutral valuation, the values of the options ($V(\epsilon,t)$, $U(\epsilon,t)$, $W(\epsilon,t)$) are determined by discounting their expected payoffs at the risk-free interest rate. By the merit of Feyman-Kac formula, for $t<T$, we can deduce the partial differential

[1]The greater the mean-reverting parameter value, $\frac{\mu}{T-t}$, the greater is the pull back to the equilibrium level

equations (PDE) form of all three options.

$$\frac{\partial H}{\partial t} + \frac{1}{2}\gamma^2 \frac{\partial^2 H}{\partial \epsilon^2} - \frac{\mu \epsilon}{T-t}\frac{\partial H}{\partial \epsilon} - rH = 0 \qquad (2.2)$$

where $H(\epsilon, t) = V(\epsilon, t), U(\epsilon, t), W(\epsilon, t)$, and r is the riskless interest rate which is assumed to be constant.

2.1.2 No Position Limits

Without taking consideration of position limits, close out a long position prior to maturity means take a simple short arbitrage position. This will yield a net benefit $-\epsilon$, however, simultaneously it costs us C_2 for early closing out of arbitrage position. Therefore, the value of $V(\epsilon, t)$ should have a lower bound of $-\epsilon - C_2$, mathematically speaking,

$$V(\epsilon, t) \geq \max(-\epsilon - C_2, 0) \qquad (2.3)$$

Similarly, close out a short position early is equivalent to take a simple long arbitrage position. This will give an profit of ϵ, however, at the same time, we will incur a cost of C_2. Therefore, the value of $U(\epsilon, t)$ should have a lower bound of $\epsilon - C_2$, mathematically speaking,

$$U(\epsilon, t) \geq \max(\epsilon - C_2, 0) \qquad (2.4)$$

Things become a little bit different to initiate a simple long or short arbitrage position. Initiating a simple long arbitrage position will yield an profit of ϵ but incur a cost of C_1. Alternatively, initiating a simple short arbitrage position will yield an profit of $-\epsilon$ but incur a cost of C_1. Sum it up, the value of $W(\epsilon, t)$ should have a lower bound of the larger value between $\epsilon + V(\epsilon, t) - C_1$ and $-\epsilon + U(\epsilon, t) - C_1$, mathematically speaking,

$$W(\epsilon, t) \geq \max(\epsilon + V(\epsilon, t) - C_1, -\epsilon + U(\epsilon, t) - C_1, 0) \qquad (2.5)$$

At maturity date, namely $t = T$, the simple arbitrage profit ϵ becomes zeros and so does these options whose underlying asset is the simple arbitrage profit. Hence

$$V(0,T) = U(0,T) = W(0,T) = 0 \tag{2.6}$$

Up till now we have derived that V, U and W follow the PDE (2.2). They are subjected to the lower bound conditions (2.3), (2.4) and (2.5). The terminal condition is (2.6).

To summarize, we solve the following problem on $(\epsilon, t) \in \{(-\infty, \infty) \times [0, T)\}$

$$\min\left\{-\frac{\partial H}{\partial t} - \frac{1}{2}\gamma^2\frac{\partial^2 H}{\partial \epsilon^2} + \frac{\mu\epsilon}{T-t}\frac{\partial H}{\partial \epsilon} + rH, H - G\right\} = 0 \tag{2.7}$$

where G is the lower bound function

$$G = \begin{cases} -\epsilon - C_2 & \text{if } H = V \\ \epsilon - C_2 & \text{if } H = U \\ \max\left(\epsilon + V, -\epsilon + U\right) - C_1 & \text{if } H = W \end{cases}$$

with the terminal condition,

$$H(\epsilon = 0, t = T) = 0$$

This variational inequality form of all three options is similar to the model of American put option P_A on $(S, t) \in \{(0, \infty) \times [0, T)\}$.

$$\min\left\{-\frac{\partial P_A}{\partial t} - \frac{1}{2}S^2\sigma^2\frac{\partial^2 P_A}{\partial S^2} - rS\frac{\partial P_A}{\partial S} + rP_A, P_A - (X - S)\right\} = 0$$

with the terminal condition,

$$P_A(S,T) = \max(X - S, 0)$$

In next subsection, we can use the same technique, projected SOR method, for implementing American put option, to implement the model (2.7) numerically.

2.1.3 With Position Limits

Next, without loss of generality, let us assume that the arbitrageur is restricted to a single net long or short arbitrage position at any moment in time. It is a reasonable assumption because of capital requirements or self-imposed exposure limits. It makes more realistic case but also adds complexity into the model.
With a position limit, closing an arbitrage position not only yields an profit but also gives the right to initiate a new arbitrage position in the future. Therefore, compared to no position limits case, the only difference in lower bound is an additional term $W(\epsilon, t)$. Hence we have

$$V(\epsilon, t) \geq \max(W(\epsilon, t) - \epsilon - C_2, 0) \tag{2.8}$$

$$U(\epsilon, t) \geq \max(W(\epsilon, t) + \epsilon - C_2, 0) \tag{2.9}$$

The value of the arbitrage option will still satisfy

$$W(\epsilon, t) \geq \max(\epsilon + V(\epsilon, t) - C_1, -\epsilon + U(\epsilon, t) - C_1, 0) \tag{2.10}$$

Of course, at maturity, $\epsilon = 0$, and all three options have no value, so that

$$V(0, T) = U(0, T) = W(0, T) = 0 \tag{2.11}$$

At this stage we have derived that V, U and W follow the PDE (2.2). They are subjected to the lower bound conditions (2.8), (2.9) and (2.10). The terminal condition is (2.11).
To summarize, we solve the following problem on $(\epsilon, t) \in \{(-\infty, \infty) \times [0, T)\}$

$$\min\left\{-\frac{\partial H}{\partial t} - \frac{1}{2}\gamma^2\frac{\partial^2 H}{\partial \epsilon^2} + \frac{\mu\epsilon}{T-t}\frac{\partial H}{\partial \epsilon} + rH, H - G\right\} = 0 \tag{2.12}$$

where G is the lower bound function

$$G = \begin{cases} W - \epsilon - C_2 & \text{if } H = V \\ W + \epsilon - C_2 & \text{if } H = U \\ \max(\epsilon + V, -\epsilon + U) - C_1 & \text{if } H = W \end{cases}$$

with the terminal condition,

$$H(\epsilon = 0, t = T) = 0$$

2.2 Numerical Scheme

In this section, we use fully implicit scheme and Crank-Nicolson scheme to discretize the models of the options. We take a transformation to make PDE look simpler and add some boundary conditions

2.2.1 Transformation

Let us recall the PDE (2.2),

$$\frac{\partial H}{\partial t} + \frac{1}{2}\gamma^2 \frac{\partial^2 H}{\partial \epsilon^2} - \frac{\mu\epsilon}{T-t}\frac{\partial H}{\partial \epsilon} - rH = 0$$

We take the transformation

$$x = (T-t)^{-\mu}\epsilon, \qquad Q(x,t) = H(\epsilon, t)$$

since

$$\begin{aligned}
\tfrac{\partial H}{\partial \epsilon} &= \tfrac{\partial Q}{\partial x}\tfrac{\partial x}{\partial \epsilon} = (T-t)^{-\mu}\tfrac{\partial Q}{\partial x} \\
\tfrac{\partial^2 H}{\partial \epsilon^2} &= (T-t)^{-\mu}\tfrac{\partial^2 Q}{\partial x^2}\tfrac{\partial x}{\partial \epsilon} = (T-t)^{-2\mu}\tfrac{\partial^2 Q}{\partial x^2} \\
\tfrac{\partial H}{\partial t} &= \tfrac{\partial Q}{\partial x}\tfrac{\partial x}{\partial t} + \tfrac{\partial Q}{\partial t} = \mu(T-t)^{-\mu-1}\epsilon\tfrac{\partial Q}{\partial x} + \tfrac{\partial Q}{\partial t}
\end{aligned}$$

substitute all terms into (2.2) and simplify, we get

$$\frac{\partial Q}{\partial t} + \frac{1}{2}\gamma^2 (T-t)^{-2\mu}\frac{\partial^2 Q}{\partial x^2} - rQ = 0 \tag{2.13}$$

After the transformation, we use $v(x,t) = V(\epsilon, t)$, $u(x,t) = U(\epsilon, t)$ and $w(x,t) = W(\epsilon, t)$. The new models are presented as follows.

For '**no position limits**' case, on the solution domain $(x,t) \in \{[x_{\min}, x_{\max}] \times [0,T)\}$

$$\min\left\{-\frac{\partial Q}{\partial t} - \frac{1}{2}\gamma^2 (T-t)^{-2\mu}\frac{\partial^2 Q}{\partial x^2} + rQ, Q - g_{\text{NP}}\right\} = 0 \tag{2.14}$$

where g_{NP} is the transformed lower bound function for 'no position limits' case,

$$g_{\text{NP}} = \begin{cases} -(T-t)^{\mu}x - C_2 & \text{if } Q = v \\ (T-t)^{\mu}x - C_2 & \text{if } Q = u \\ \max\left((T-t)^{\mu}x + v, -(T-t)^{\mu}x + u\right) - C_1 & \text{if } Q = w \end{cases}$$

with transformed terminal condition,

$$Q(x = 0, t = T) = 0$$

and with transformed boundary conditions,

$$Q(x_{\min}, t) = \begin{cases} -(T-t)^{\mu}x_{\min} - C_2 & \text{if } Q = v \\ 0 & \text{if } Q = u \\ -(T-t)^{\mu}x_{\min} - C_1 & \text{if } Q = w \end{cases}$$

and

$$Q(x_{\max}, t) = \begin{cases} 0 & \text{if } Q = v \\ (T-t)^{\mu}x_{\max} - C_2 & \text{if } Q = u \\ (T-t)^{\mu}x_{\max} - C_1 & \text{if } Q = w \end{cases}$$

For '**with position limits**' case, on the solution domain $(x,t) \in \{[x_{\min}, x_{\max}] \times [0,T)\}$

$$\min\left\{-\frac{\partial Q}{\partial t} - \frac{1}{2}\gamma^2(T-t)^{-2\mu}\frac{\partial^2 Q}{\partial x^2} + rQ, Q - g_{\text{WP}}\right\} = 0 \qquad (2.15)$$

where g_{WP} is the transformed lower bound function for 'with position limits' case,

$$g_{\text{WP}} = \begin{cases} w - (T-t)^{\mu}x - C_2 & \text{if } Q = v \\ w + (T-t)^{\mu}x - C_2 & \text{if } Q = u \\ \max\left((T-t)^{\mu}x + v, -(T-t)^{\mu}x + u\right) - C_1 & \text{if } Q = w \end{cases}$$

with transformed terminal condition,

$$Q(x = 0, t = T) = 0$$

and with transformed boundary conditions,

$$Q(x_{\min}, t) = \begin{cases} -2(T-t)^{\mu} x_{\min} - C_1 - C_2 & \text{if } Q = v \\ 0 & \text{if } Q = u \\ -(T-t)^{\mu} x_{\min} - C_1 & \text{if } Q = w \end{cases}$$

and

$$Q(x_{\max}, t) = \begin{cases} 0 & \text{if } Q = v \\ 2(T-t)^{\mu} x_{\max} - C_1 - C_2 & \text{if } Q = u \\ (T-t)^{\mu} x_{\max} - C_1 & \text{if } Q = w \end{cases}$$

For '*no position limits*' case, the system of PDEs (2.14) is easy to solve because they are not really 'coupled'. We can solve the first two variational equations on their own, just similar to deal with the American options, then using the results solved by first two variational equations to solve the third variational equation. The three options values do not need to be solved simultaneously.
For '*with position limits*' case, the system of PDEs (2.15) is nested, the variational inequality of each option involves the value of at least one other options. We need to solve these options simultaneously at each time step. We adopt an iterative method, and stop the iteration when the value of each option changes is within a preset tolerance in two consecutive iterations.

2.2.2 Numerical Discretization

The solution region is confined as

$$\Omega = \{(x,t)\,|x_{\min} \le x \le x_{\max}, 0 \le t \le T\} \tag{2.16}$$

The grid for the finite difference scheme is defined as followed:

$$x_i = x_{\min} + i \cdot \delta x, \quad i = 0, 1, \cdots, m, \quad x_0 = x_{\min},\ x_m = x_{\max}$$

$$t_j = j \cdot \delta t, \quad j = 0, 1, \cdots, n, \quad t_0 = 0,\ t_n = T$$

where

$$\delta x = \frac{x_{\max} - x_{\min}}{m}, \quad \delta t = \frac{T}{n}$$

Define the grid function

$$Q = \{Q_{i,j} \,|0 \leq i \leq m, 0 \leq j \leq n\} \tag{2.17}$$

where

$$Q_{i,j} := Q(x_i, t_j) \text{ for } 0 \leq i \leq m, 0 \leq j \leq n$$

Equation (2.13) can be discretized by a standard one factor finite difference method with variable timeweighting to give

$$\begin{aligned} Q_{i,j+1} - Q_{i,j} = (1-\theta)\left[-\alpha_{j+1}Q_{i+1,j+1} - \beta_{j+1}Q_{i,j+1} - \alpha_{j+1}Q_{i-1,j+1}\right] \\ +\theta\left[-\alpha_j Q_{i+1,j} - \beta_j Q_{i,j} - \alpha_j Q_{i-1,j}\right] \end{aligned} \tag{2.18}$$

$\theta = 1$ for fully implicit scheme, and $\theta = 0.5$ for Crank-Nicolson scheme.
For notational convenience, it helps to rewrite the above discrete equations in matrix form. Let

$$\mathbf{Q}_{j+1} = [Q_{1,j+1}, Q_{2,j+1}, \cdots, Q_{m-1,j+1}]^T \qquad \mathbf{Q}_j = [Q_{1,j}, Q_{2,j}, \cdots, Q_{m-1,j}]^T$$

and we obtain a compact matrix form

$$(\mathbf{I} + \theta\mathbf{M})\,\mathbf{Q}_j = (\mathbf{I} - (1-\theta)\mathbf{M})\,\mathbf{Q}_{j+1} + \mathbf{b} \tag{2.19}$$

where matrix $\mathbf{I}$ is an identical matrix, vector $\mathbf{b}$ handles the boundary conditions, and tri-diagonal matrix $\mathbf{M}$ is

$$\mathbf{M} = -\begin{bmatrix} \beta_1 & \alpha_1 & & & \\ \alpha_2 & \beta_2 & \alpha_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \alpha_{m-2} & \beta_{m-2} & \alpha_{m-2} \\ & & & \alpha_{m-1} & \beta_{m-1} \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} \theta\alpha_1 Q_{0,j} + (1-\theta)\alpha_1 Q_{0,j+1} \\ 0 \\ \vdots \\ 0 \\ \theta\alpha_{m-1}Q_{m,j} + (1-\theta)\alpha_{m-1}Q_{m,j+1} \end{bmatrix}$$

where $\alpha_i = \frac{\delta t}{2\delta x^2}\gamma^2(T-t_j)^{-2\mu}$ and $\beta_i = -2\alpha_i - r\delta t$.

The matrix $\mathbf{I}+\theta\mathbf{M}$ is a row diagonally dominant matrix, hence the projected SOR ensures the convergence of the numerical solutions. The overrelaxtion method should take into account the tri-diagonal nature of the matrix $\mathbf{I}+\theta\mathbf{M}$, and it should also be adjusted for early exercise. Let $g_{i,j}$, $i = 1, 2, \cdots, m-1$, be the intrinsic value when $x = x_i$. Therefore,

$$g_{i,j} = \begin{cases} \max\left\{-x_i(T-t_j)^\mu - C_2, 0\right\} & \text{if } Q = v \\ \max\left\{x_i(T-t_j)^\mu - C_2, 0\right\} & \text{if } Q = u \\ \max\left\{\max(x_i(T-t_j)^\mu + v_{i,j}, -x_i(T-t_j)^\mu + u_{i,j}) - C_1, 0\right\} & \text{if } Q = w \end{cases}$$

and we denote the right hand of equation (2.19),

$$\mathbf{z}_{j+1} = \left(\mathbf{I} - (1-\theta)\mathbf{M}\right)\mathbf{Q}_{j+1} + \mathbf{b}$$

For each time layer j, let $\mathbf{Q}_j^k$ be the kth estimate for $\mathbf{Q}_j$, the projected SOR method for '*no position limits*' case can then be written as in Algorithm 1.

Algorithm 1 Pseudo Code of Projected SOR Method for No Position Limits, One Dimensional Problem. Determining Option Values $Q_{i,j}$ for Interior Node (x_i, t_j)

Let $\mathbf{Q}_j^0 = \mathbf{Q}_{j+1}$

for $k = 0, 1, 2\cdots$ until convergence **do**

  if $i = 1$

    $Q_{i,j}^{k+1} = \max\left\{g_{i,j}, Q_{i,j}^k + \frac{\omega}{1-\theta\beta_i}\left[z_{i,j+1} - (1-\theta\beta_i)Q_{i,j}^k + \theta\alpha_i Q_{i+1,j}^k\right]\right\}$

  elseif $i = 2 : m-2$

    $Q_{i,j}^{k+1} = \max\left\{g_{i,j}, Q_{i,j}^k + \frac{\omega}{1-\theta\beta_i}\left[z_{i,j+1} + \theta\alpha_i Q_{i-1,j}^{k+1} - (1-\theta\beta_i)Q_{i,j}^k + \theta\alpha_i Q_{i+1,j}^k\right]\right\}$

  elseif $i = m-1$

    $Q_{i,j}^{k+1} = \max\left\{g_{i,j}, Q_{i,j}^k + \frac{\omega}{1-\theta\beta_i}\left[z_{i,j+1} + \theta\alpha_i Q_{i-1,j}^{k+1} - (1-\theta\beta_i)Q_{i,j}^k\right]\right\}$

  end if

  if $\left\|\mathbf{Q}_j^{k+1} - \mathbf{Q}_j^k\right\| < tolerance$ **then**

    Quit the iterations

  end if

end for

In Algorithm 1, we solve v and u independently, and use the results to solve w finally. It is not a very difficult task, however, for '*with position limits*' case, we need to solve v, u and w at the same time. The intrinsic values for them are[2]

$$\begin{cases} g^v_{i,j} = \max\{w_{i,j} - x_i(T-t_j)^\mu - C_2, 0\} \\ g^u_{i,j} = \max\{w_{i,j} + x_i(T-t_j)^\mu - C_2, 0\} \\ g^w_{i,j} = \max\{\max(x_i(T-t_j)^\mu + v_{i,j}, -x_i(T-t_j)^\mu + u_{i,j}) - C_1, 0\} \end{cases}$$

For each time layer j, let $\mathbf{v}^k_j$, $\mathbf{u}^k_j$ and $\mathbf{w}^k_j$ be the kth estimate for $\mathbf{v}_j$, $\mathbf{u}_j$ and $\mathbf{w}_j$. We present the projected SOR method for *with position limits* case which can then be written as in Algorithm 2.

Algorithm 2 Pseudo Code of Projected SOR Method for With Position Limits, One Dimensional Problem. Determining Option Values $Q_{i,j}$ for Interior Node (x_i, t_j)

Let $\mathbf{v}^0_j = \mathbf{v}_{j+1}$, $\mathbf{u}^0_j = \mathbf{u}_{j+1}$ and $\mathbf{w}^0_j = \mathbf{w}_{j+1}$
for $i = 1 : m-1$
  $g^v_{i,j} = \max\{w_{i,j+1} - x_i(T-t_j)^\mu - C_2, 0\}$
  $g^u_{i,j} = \max\{w_{i,j+1} + x_i(T-t_j)^\mu - C_2, 0\}$
end for
for $k = 0, 1, 2\cdots$ until convergence **do**
  By Algorithm 1: Calculate $v^{k+1}_{i,j}$ and $u^{k+1}_{i,j}$
  for $i = 1 : m-1$
    $g^w_{i,j} = \max\left\{\max(x_i(T-t_j)^\mu + v^{k+1}_{i,j}, -x_i(T-t_j)^\mu + u^{k+1}_{i,j}) - C_1, 0\right\}$
  end for
  By Algorithm 1: Calculate $w^{k+1}_{i,j}$
  for $i = 1 : m-1$
    $g^v_{i,j} = \max\left\{w^{k+1}_{i,j} - x_i(T-t_j)^\mu - C_2, 0\right\}$
    $g^u_{i,j} = \max\left\{w^{k+1}_{i,j} + x_i(T-t_j)^\mu - C_2, 0\right\}$
  end for
  if $\left\|\left(\mathbf{v}^{k+1}_j - \mathbf{v}^k_j\right) + \left(\mathbf{u}^{k+1}_j - \mathbf{u}^k_j\right) + \left(\mathbf{w}^{k+1}_j - \mathbf{w}^k_j\right)\right\| < tolerance$ **then**
    Quit the iterations
  end if
end for

[2]We use the superscript to denote the corresponding intrinsic value

2.3 Experimental Results

In this section, we take some data inputs, calculate the options price and draw their respective exercise regions and exercise boundaries.

2.3.1 Data Inputs

The value of the simple arbitrage opportunity is defined by

$$\epsilon_t = F_t(T)e^{-r(T-t)} + \mathrm{PV}_t(\mathrm{div}) - S_t \tag{2.20}$$

where $F_t(T)$ is the futures price at time t for a contract maturing at time T, r is the riskless interest rate, $\mathrm{PV}_t(\mathrm{div})$ is the present value of the daily dividends on the S&P 500 index portfolio up to the maturity of the contract, and S_t is the value of the index at time t. We partition $N_x = 400$ and $N_t = 400$ in state and time

Input Parameter	
Rate of Mean Reversion μ	0.03
Standard Deviation γ	0.6
Riskless Interest Rate r	0.07
Time to Maturity T	1
Type One Cost C_1	1.2
Type Two Cost C_2	0.5

Table 2.1: *Model Parameters for Stylized One Dimensional Problem*

variables, and we choose Crank-Nicolson scheme for numerical experiment.

2.3.2 Option Values

We present the option values of V ,U and W, without and with position limits.

ϵ	-1.5	-1	-0.5	0	0.5	1	1.5
V	1.0000	0.5299	0.2113	0.0576	0.0099	0.0010	0.0001
U	0.0001	0.0010	0.0099	0.0576	0.2113	0.5299	1.0000
W	0.3685	0.1208	0.0258	0.0065	0.0258	0.1208	0.3685

Table 2.2: *Values of Early Close-Out and Open Options, No Position Limits*

ϵ	-1.5	-1	-0.5	0	0.5	1	1.5
V	1.3685	0.6506	0.2369	0.0609	0.0100	0.0010	0.0001
U	0.0001	0.0010	0.0101	0.0609	0.2369	0.6506	1.3685
W	0.3685	0.1208	0.0258	0.0065	0.0258	0.1208	0.3685

Table 2.3: *Values of Early Close-Out and Open Options, With Position Limits*

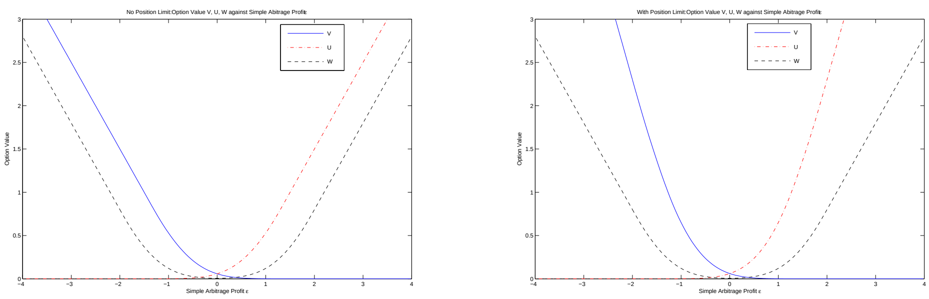

Figure 2.1: *The Value of Three Options, Without and With Position Limits*

1. **Value of V and U are larger for '*with position limits*' case**

   List out the variational equations for both cases.

   For '*no position limits*' case

$$\min\left\{-\frac{\partial V}{\partial t}-\frac{1}{2}\gamma^2\frac{\partial^2 V}{\partial\epsilon^2}+\frac{\mu\epsilon}{T-t}\frac{\partial V}{\partial\epsilon}+rV, V-(-\epsilon-C_2)\right\}=0$$
$$\min\left\{-\frac{\partial U}{\partial t}-\frac{1}{2}\gamma^2\frac{\partial^2 U}{\partial\epsilon^2}+\frac{\mu\epsilon}{T-t}\frac{\partial U}{\partial\epsilon}+rU, U-(\epsilon-C_2)\right\}=0$$

For '*with position limits*' case

$$\min\left\{-\tfrac{\partial V}{\partial t}-\tfrac{1}{2}\gamma^2\tfrac{\partial^2 V}{\partial \epsilon^2}+\tfrac{\mu\epsilon}{T-t}\tfrac{\partial V}{\partial \epsilon}+rV, V-(W-\epsilon-C_2)\right\}=0$$
$$\min\left\{-\tfrac{\partial U}{\partial t}-\tfrac{1}{2}\gamma^2\tfrac{\partial^2 U}{\partial \epsilon^2}+\tfrac{\mu\epsilon}{T-t}\tfrac{\partial U}{\partial \epsilon}+rU, U-(W+\epsilon-C_2)\right\}=0$$

Clearly, the options of V and U for 'with position limits' case have an additional non-negative value in the lower bound.

2. **Value of W is exactly the same for both cases**

   In both cases we have

$$\min\left\{-\frac{\partial W}{\partial t}-\frac{1}{2}\gamma^2\frac{\partial^2 W}{\partial \epsilon^2}+\frac{\mu\epsilon}{T-t}\frac{\partial W}{\partial \epsilon}+rW, W-\max\left(\epsilon+V-C_1,-\epsilon+U-C_1\right)\right\}=0$$

   When early exercise happens for W, W can either takes the value of $\epsilon+V-C_1$ or $-\epsilon+U-C_1$. When W takes $\epsilon+V-C_1$, which means ϵ is positive in large, hence V goes to zero. In another hand, when W takes $-\epsilon+U-C_1$, which means ϵ is negative in large, hence U approaches to zero. So effectively, the variational inequality of W reduced to

$$\min\left\{-\frac{\partial W}{\partial t}-\frac{1}{2}\gamma^2\frac{\partial^2 W}{\partial \epsilon^2}+\frac{\mu\epsilon}{T-t}\frac{\partial W}{\partial \epsilon}+rW, W-\max\left(\epsilon-C_1,-\epsilon-C_1\right)\right\}=0$$

   The variational equation above indicates that W is independent of values of V and U. Therefore, for both models, W are identical although V and U have different values. Economically speaking, it means the value of the option for investor to initiate an arbitrage position is not affected by existence of position limits.

2.3.3 Exercise Region and Boundary

For '*no position limits*' case: According to the original variational inequality for V

$$\min\left\{-\frac{\partial V}{\partial t}-\frac{1}{2}\gamma^2\frac{\partial^2 V}{\partial \epsilon^2}+\frac{\mu\epsilon}{T-t}\frac{\partial V}{\partial \epsilon}+rV, V-(-\epsilon-C_2)\right\}=0$$

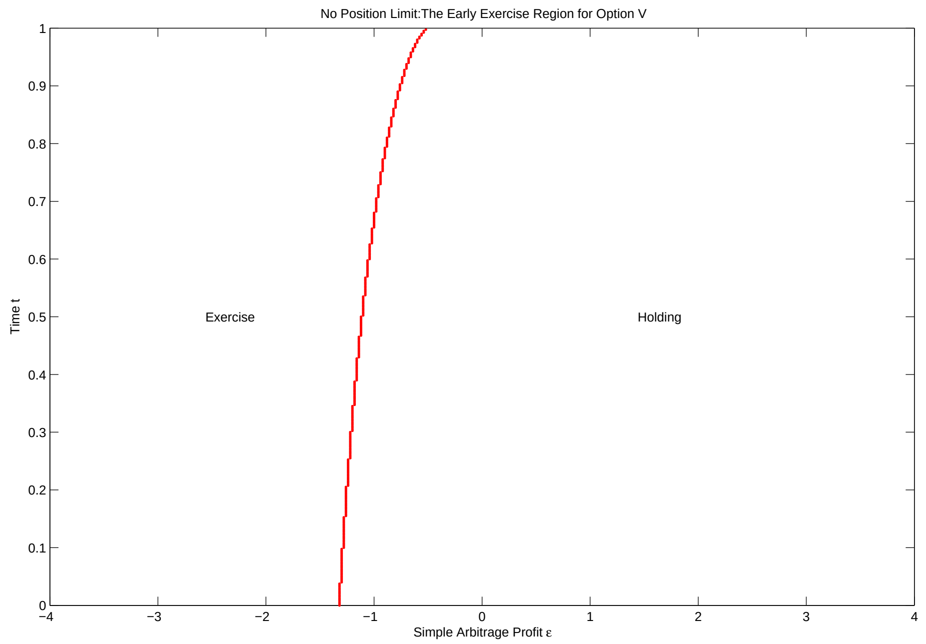

Figure 2.2: *For No Position Limits Case: The Early Exercise Region of Option V*

We would expect early exercise to occur only when $-\epsilon - C_2 \geq 0$, namely, when ϵ is negative and $|\epsilon|$ is large. From Figure 2.2, we can see that the exercise region is below $\epsilon = -C_2$, which agrees with our expectation. Furthermore, a closer look shows us that the exercise boundary is monotonically increasing, which shows that the closer to maturity we are, the smaller $|\epsilon|$ value is required for early exercise to occur. According to the original variational inequality for U

$$\min\left\{-\frac{\partial U}{\partial t} - \frac{1}{2}\gamma^2\frac{\partial^2 U}{\partial \epsilon^2} + \frac{\mu\epsilon}{T-t}\frac{\partial U}{\partial \epsilon} + rU, U - (\epsilon - C_2)\right\} = 0$$

We would expect early exercise to occur only when $\epsilon - C_2 \geq 0$, namely, when ϵ is positive and large. From Figure 2.3, we can see that the exercise region is above $\epsilon = C_2$, which agrees with our expectation. Furthermore, a closer look shows us that the exercise boundary is monotonically decreasing, which shows that the closer to maturity we are, the smaller ϵ value is required for early exercise to occur.

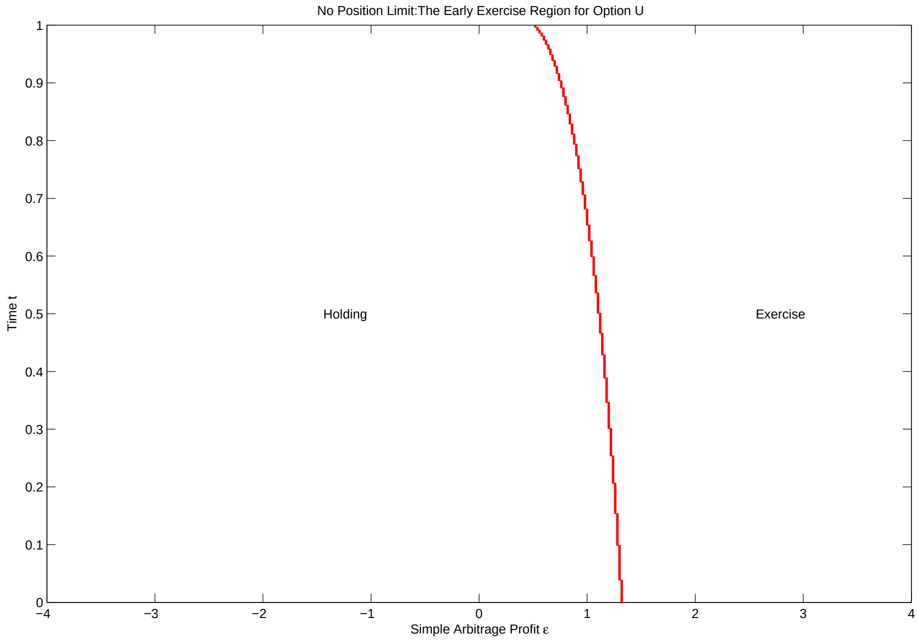

Figure 2.3: *For No Position Limits Case: The Early Exercise Region of Option U*

According to the original variational inequality for W

$$\min\left\{-\frac{\partial W}{\partial t}-\frac{1}{2}\gamma^2\frac{\partial^2 W}{\partial\epsilon^2}+\frac{\mu\epsilon}{T-t}\frac{\partial W}{\partial\epsilon}+rW, W-\max\left(\epsilon+V-C_1,-\epsilon+U-C_1\right)\right\}=0$$

We would expect early exercise to occur only when $\max\left(\epsilon+V-C_1,-\epsilon+U-C_1\right)\geq 0$. So we would expect early exercise to occur when ϵ is eith positive or negative in large. From Figure 2.4, we can see that the exercise region is above $\epsilon=C_1$ or below $\epsilon=-C_1$, which agrees with our expectation. Furthermore, a closer look shows us that the exercise boundary is monotonically approaching $\epsilon=C_1$ for the upper early exercise region, and monotonically approaching $\epsilon=-C_1$ for the lower early exercise region. This shows that the closer to maturity we are, the smaller $|\epsilon|$ is required for early exercise to occur. Figure 2.5 summarizes the early exercise boundaries for all three options. By our model, the options should be exercised, namely, the arbitrage positions should be closed out or initiated once ϵ reaches the

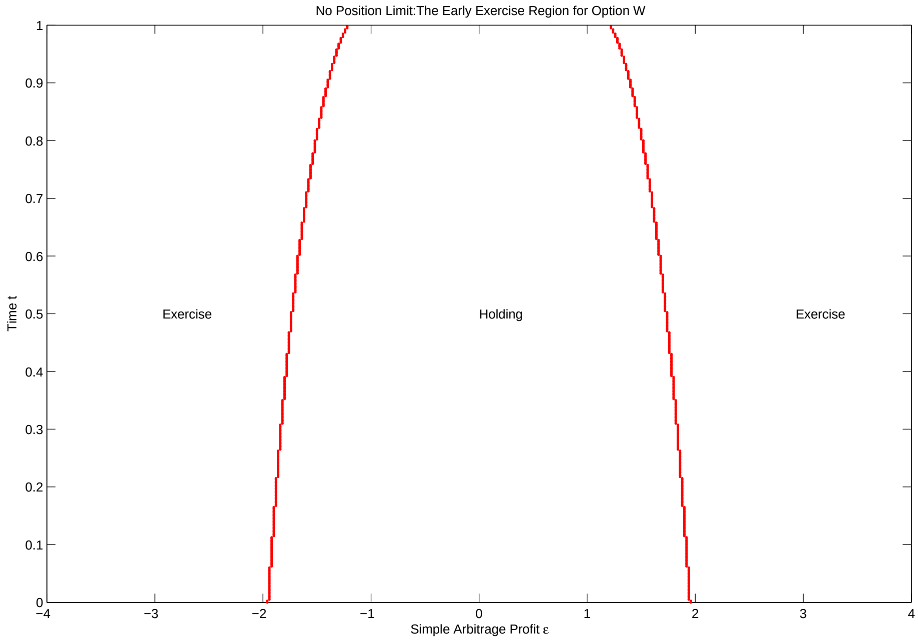

Figure 2.4: *For No Position Limits Case: The Early Exercise Region of Option W*

boundaries at a certain time t.

Interestingly, for '*with position limits*' case, the exercise regions and boundaries for exactly the same with the '*no position limits*' case. This implies that whether an investor is subjected to position limits or not, she should adopt the same optimal arbitrage strategy.

2.3.4 Effects of changing input values

Varying the input parameters of the program will produce different pattern of early exercise region and option values. We have six input parameters, namely rate of mean reversion μ, standard deviation γ, riskless interest rate r, time to maturity T, type one cost C_1 and type two cost C_2. We are particularly interested in μ, γ and T. When we vary one single input parameter to simulate the exercise region

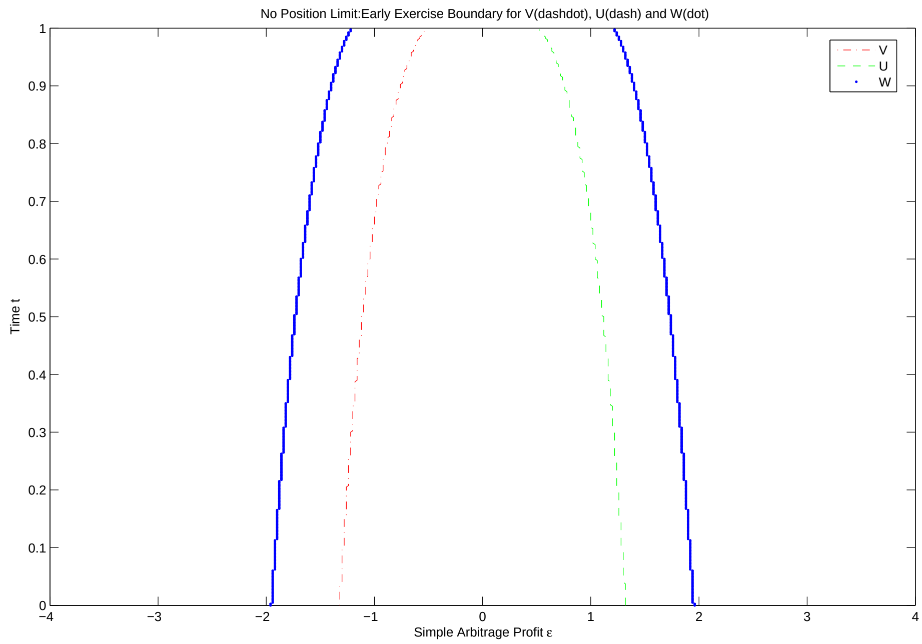

Figure 2.5: *For No Position Limits Case: The Early Exercise Boundary of Three Options*

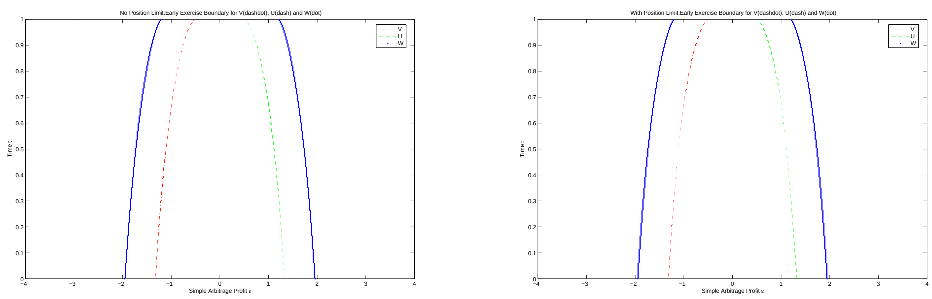

Figure 2.6: *For Both Cases: The Early Exercise Boundaries of Three Options*

and option values, we hold other parameters unchanged. In this section, we first use Monte Carlo simulation to generate the path of ϵ given different values of μ, γ and T, in the purpose of observing the effects on the realization of simple arbitrage profit. Then, we display the plots of option values and early exercise boundaries, and compare the plots for different values of μ, γ and T. Because the early exercise boundaries are exactly the same under two cases, we only show the plot under 'no position limits' case for illustration.

Change rate of mean reversion μ

Figure 2.7 report the path of simple arbitrage profit with step size 500 and 20. The first plot is extremely messy and difficult to see the pattern, and we use the second one for illustration purposes. In Figure 2.7, the path drawn in solid line (cyan) is always above path drawn in dashdot line (black) and dash line (red). Therefore we conjecture that ϵ approaches to its mean level 0 faster with a higher value of μ. It makes perfect sense cause μ measure the speed to its mean level. For all three options, namely, option to close a long (short) arbitrage position

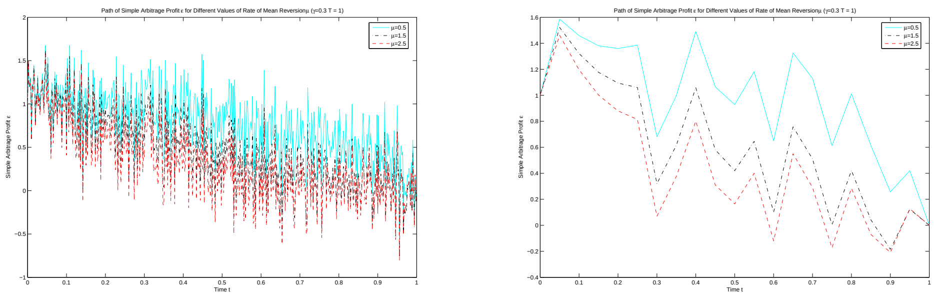

Figure 2.7: *The Path of Simple Arbitrage Profit ϵ with Different μ, $N = 500, 20$*

and option to initiate arbitrage position, their values are negatively related to μ. The reason is higher μ brings ϵ to zero more quickly, which decreases the value of three options V, U and W. Figure 2.8 displays the relationship between option values and mean reversion μ. The solid line (cyan), dashdot line (black) and dash line (red) correspond the case when $\mu = 0.5, 1.5, 2.5$ respectively. In the plot of early exercise boundaries, Figure 2.9, the option V, U and W are labeled. For these three options' exercise boundaries, when μ is decreasing, thse boundaries

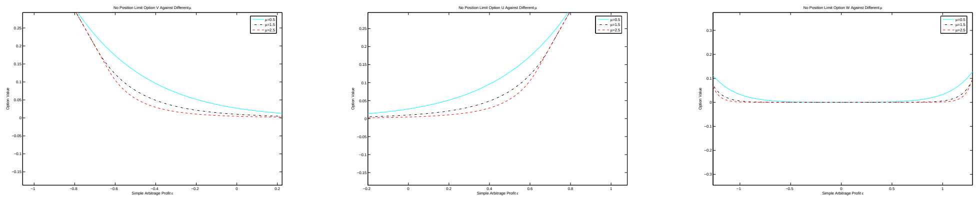

Figure 2.8: *The Option Values with Different Mean Reversion μ*

are spreading out (far away from $\epsilon = 0$). We take the option V (solid line) for illustration. At initial time, a largest negative value of ϵ is required for early exercise for a smallest μ, because of the lowest possibility of dragging ϵ to zero which makes option worthless. As a result, we can actually hold the option V until the ϵ become quite negative large. (This is attractive for exercising option V). Moreover, the three exercise boundaries converge to one point at maturity while the option value is zero regardless of the value of μ.

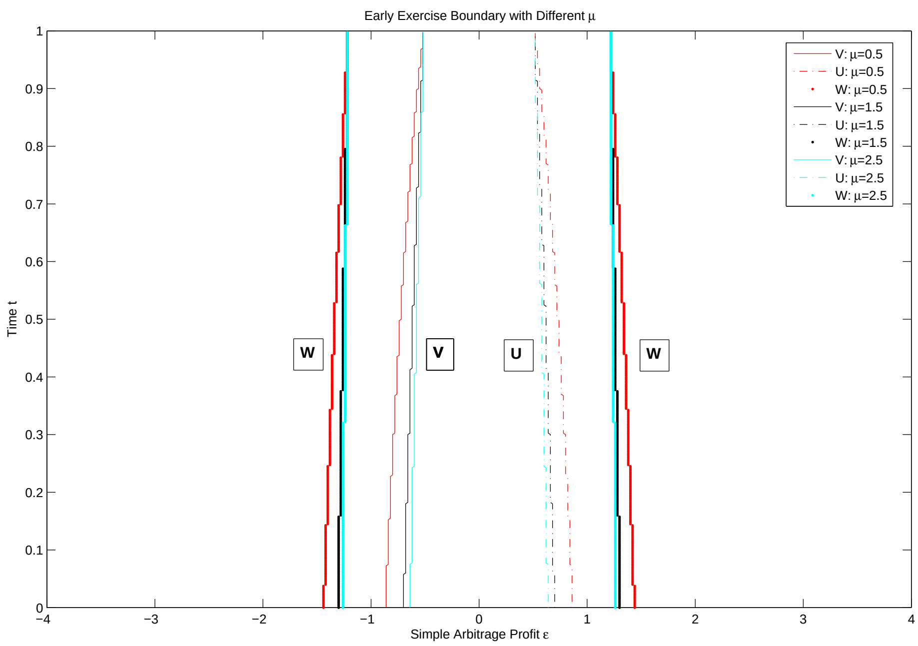

Figure 2.9: *The Early Exercise Boundaries with Different Mean Reversion μ*

Change standard deviation γ

In Figure 2.10, the path drawn in dash line (red) fluctuates in a larger amplitude than the path drawn in dashdot line (black) and in solid line (cyan) do, hence, simple arbitrage profit ϵ, analogous as stock S in standard option, has a higher chance to reach larger and smaller values, which increases the option values. Volatility γ for this type option plays the similar role of volatility σ for standard option. The option price is monotonously increasing with respect to volatility γ. Figure 2.11 shows the relationship between option values and volatility γ. The solid line (cyan), dashdot line (black) and dash line (red) correspond the case when $\gamma = 0.3, 0.6, 0.9$ respectively. In these three subplots, the dash (red) line (option value curve with largest γ) is clearly above the dashdot (black) and solid (cyan) line (option value curve with smaller γ). This observation meets our expectation. In the plot of

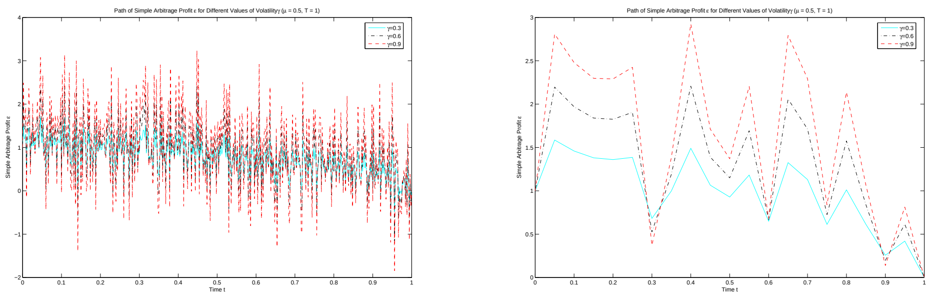

Figure 2.10: *The Path of Simple Arbitrage Profit* ϵ *with Different* γ, $N = 500, 20$

early exercise boundaries, Figure 2.12, the option V, U and W are labeled. For these three options' exercise boundaries, when γ is increasing, thse boundaries are spreading out (far away from $\epsilon = 0$). We take the option V (solid line) for illustration. At initial time, a smallest negative value of ϵ is required for early exercise because investors bet ϵ will be more likely go far away from zero with largest value of γ. It is interesting to find that the boundaries are further away from the line $\epsilon = 0$ when γ increases. It means that with larger γ value, larger absolute value of ϵ

is required for early exercise of the options to occur. The reason is higher volatility of ϵ makes investor more confident to wait until large absolute value of ϵ to realize before taking any actions. While time approaches maturity, investor's confidence about volatile ϵ is dampened. Therefore, the early exercise boundaries converge ultimately. When time approaches to maturity, large volatility is not alluring for investors to keep option unexercised, hoping for big movement in arbitrage profit, because of insufficient time left for them to make decision.

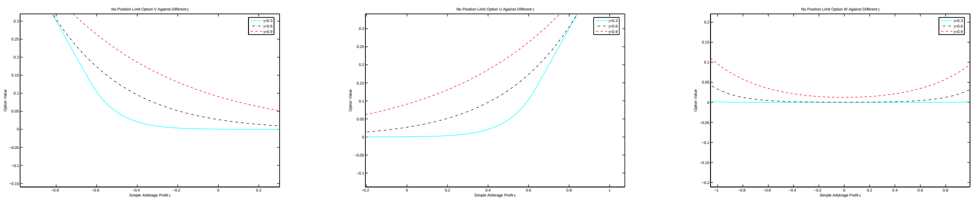

Figure 2.11: *The Option Values with Different Volatility* γ

Change time to maturity T

Intuitively speaking, option has more values with longer maturity because the option holders have more freedom to exercise it early. Figure 2.14 reports how option values are related to maturity T. The solid line (cyan), dashdot line (black) and dash line (red) correspond the case when $T = 1, 2, 3$ respectively. The dash line (red), represented by the largest option value, lies topmost. Figure 2.15 displays the early exercise boundaries with different maturity T. The option V, U and W are labeled. For these three options' exercise boundaries, when T is increasing, thse boundaries are spreading out (far away from $\epsilon = 0$). At initial time, a largest absolute value of ϵ is required for early exercise for a longest T, since investor have more confidence on those options with longer maturity.

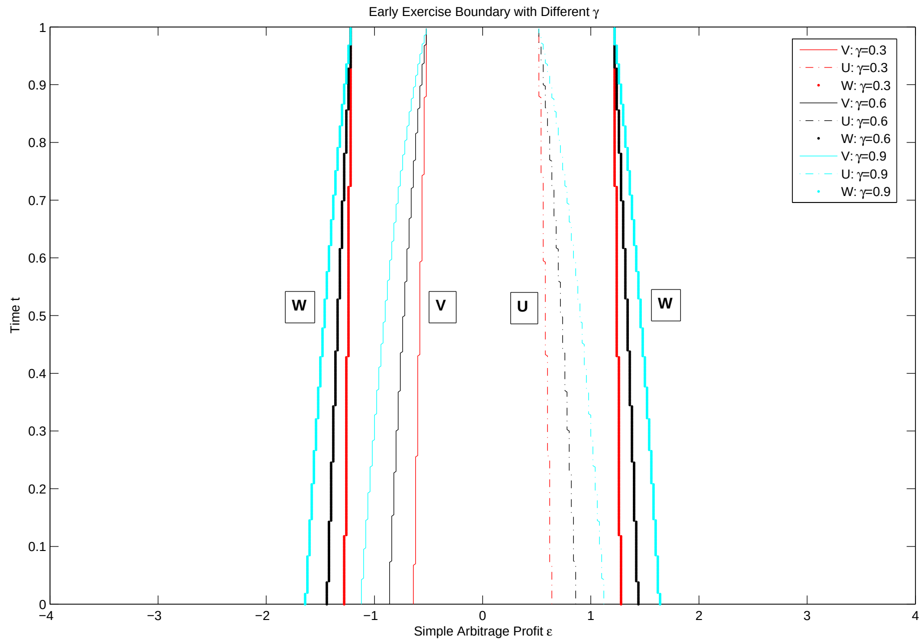

Figure 2.12: *The Early Exercise Boundaries with Different Volatility γ*

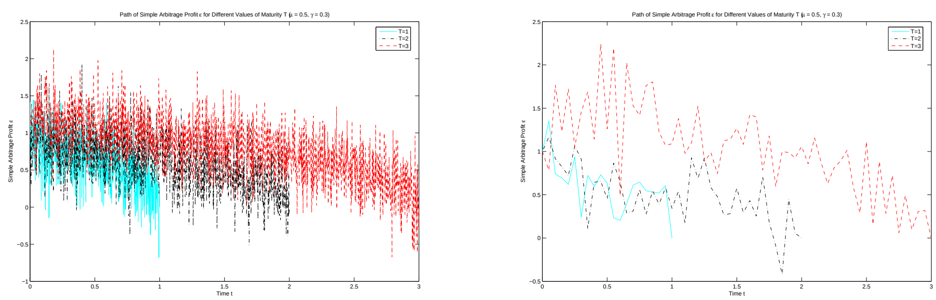

Figure 2.13: *The Path of Simple Arbitrage Profit ϵ with Different T, $N = 500, 20$*

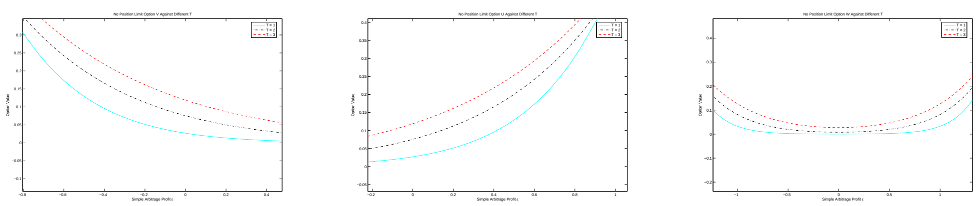

Figure 2.14: *The Option Values with Different Maturity T*

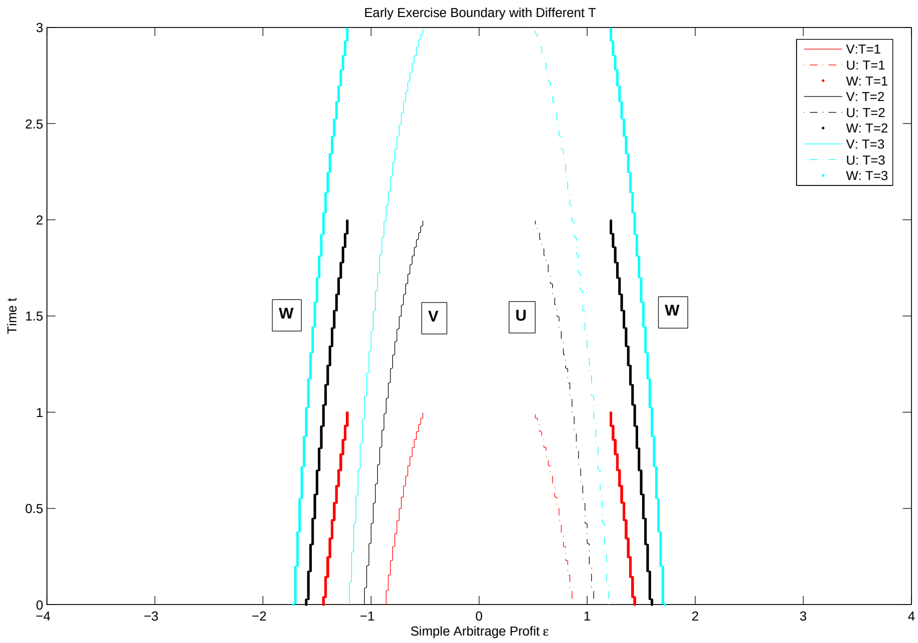

Figure 2.15: *The Early Exercise Boundaries with Different Maturity T*

Chapter 3

Two Dimensional Problem

3.1 Theoretical Model

In this section we focus on two dimensional problem and derive the partial differential equation for the options to close out or initiate a stock index arbitrage position, and construct the complete model for ‘no position limits’ case and ‘with position limits’ case.

3.1.1 Order Imbalance

In principle, the value of these options $V(\cdot)$, $U(\cdot)$ and $W(\cdot)$, may depend on additional state variables. Recall the SDE of underlying asset ϵ

$$d\epsilon = -\frac{\mu\epsilon}{T-t}dt + \gamma dW$$

In one dimensional problem, we treat μ as constant. From financial point of view, μ is the rate of mean reversion which measures the speed that ϵ approaches its mean-reversion level 0. In [3], Brennan and Schartz suggested that in particular for days that are far away from maturity, the critical ϵ values are sensitive to the parameter estimates and mentioned that we could reject the constancy of the mean reversion parameter across contracts. To test the robustness of an economic model, we want to verify whether a model can capture the main economic phenomena even when some stochastic model parameters are held deterministic. In this section, we modify the mean reversion coefficient μ to make it stochastic and examine the impact on the critical ϵ values.

Hence, we introduce a new state variable, named order imbalance and denoted by I. Order imbalance has also been found to have a significant impact on stock returns. It is defined as the difference between the dollar volume crossed at ask prices and that crossed at bid prices. Trades executed at ask prices represent buyer-initiated trades and those executed at bid prices represent seller-initiated trades. A positive order imbalance indicates that buying is more active than selling, whereas a negative order imbalance indicates that selling is more active than buying. The variable $I = 0$, indicating a balance in actual order book in stock index futures market, means there is no simple arbitrage profit existing in the market. Mathematically, ϵ is dragged to its mean level 0 very quickly, indicated by a large value of μ. Fung (2007) pointed out in [4] that on average, positive order imbalance is associated with positive arbitrage basis and negative order imbalance is associated with negative arbitrage basis. Therefore, we model the mean reversion coefficient μ as follows.

$$\mu = c + d \cdot \mathrm{sgn}(\epsilon) \cdot I \tag{3.1}$$

where c is assumed to be constant and d is a positive constant to make μ positive. The mathematical modeling for μ is intuitively correct.

- When positive I increases, there are more long positions than short positions on futures, which would decrease the value of ϵ. So, μ should be an increasing function of I when ϵ is positive, and a decreasing function of I when ϵ is negative.

$$\mu = \begin{cases} c + dI & \text{when } \epsilon > 0 \quad \text{increasing w.r.t } I \\ c - dI & \text{when } \epsilon < 0 \quad \text{decreasing w.r.t } I \end{cases}$$

- When negative I decreases, there are more short positions than long positions on futures, which would increase the value of ϵ. So, μ should be a decreasing function of I when ϵ is positive, and an increasing function of I when ϵ is negative.

$$\mu = \begin{cases} c - d(-I) & \text{when } \epsilon > 0 \quad \text{decreasing w.r.t } -I \\ c + d(-I) & \text{when } \epsilon < 0 \quad \text{increasing w.r.t } -I \end{cases}$$

Now we have two SDEs for state variables

$$\begin{cases} d\epsilon = -\frac{\mu\epsilon}{T-t}dt + \gamma dW_1 \\ dI = -aIdt + bdW_2 \end{cases} \tag{3.2}$$

where a and b are constants and W_1 and W_2 are correlated with correlation coefficient ρ. For the second SDE, we add mean reverting to model stochastic process of I.

Let $V(\epsilon, I, t)(U(\epsilon, I, t))$ be the value of the right to close a long (short) arbitrage position prior to maturity when the simple arbitrage profit before transaction costs is ϵ and the time to maturity of the futures contract is $T-t$. Similarly, let $W(\epsilon, I, t)$ be the value of the right to initiate an arbitrage position.

By risk neutral valuation, the values of the options ($V(\epsilon, I, t)$, $U(\epsilon, I, t)$, $W(\epsilon, I, t)$) are determined by discounting their expected payoffs at the risk-free interest rate. By the merit of Feyman-Kac formula, for $t < T$, we can deduce the partial differential equations (PDE) form of all three options.

$$\frac{\partial H}{\partial t} + \frac{1}{2}\gamma^2\frac{\partial^2 H}{\partial \epsilon^2} + \rho\gamma b\frac{\partial^2 H}{\partial \epsilon \partial I} + \frac{1}{2}b^2\frac{\partial^2 H}{\partial I^2} - \frac{(c + d \cdot \text{sgn}(\epsilon)I)\epsilon}{(T-t)}\frac{\partial H}{\partial \epsilon} - aI\frac{\partial H}{\partial I} - rH = 0 \tag{3.3}$$

where $H(\epsilon, I, t) = V(\epsilon, I, t), U(\epsilon, I, t), W(\epsilon, I, t)$, and r is the riskless interest rate which is assumed to be constant.

Denote $\mathcal{L}$ as linear operator

$$\mathcal{L} = \frac{\partial}{\partial t} + \frac{1}{2}\gamma^2\frac{\partial^2}{\partial \epsilon^2} + \rho\gamma b\frac{\partial^2}{\partial \epsilon \partial I} + \frac{1}{2}b^2\frac{\partial^2}{\partial I^2} - \frac{(c + d \cdot \mathrm{sgn}(\epsilon)I)\epsilon}{(T-t)}\frac{\partial}{\partial \epsilon} - aI\frac{\partial}{\partial I} - r$$

We will use this to simplify the lengthy operator in the next sections.

3.1.2 No Position Limits

Under 'no position limits' assumption, closing out a long (short) position prior to maturity means take a simple short (long) arbitrage position. Therefore, we can identify the lower bounds for three options.

$$\begin{cases} V(\epsilon, I, t) \geq \max\{-\epsilon - C_2, 0\} \\ U(\epsilon, I, t) \geq \max\{\epsilon - C_2, 0\} \\ W(\epsilon, I, t) \geq \max\{\epsilon + V(\epsilon, I, t) - C_1, -\epsilon + U(\epsilon, I, t) - C_1, 0\} \end{cases}$$

For two dimensional problem, the difficult part is to identify its boundary condition. Take $V(\epsilon, I, t)$, the option to close a long arbitrage position, as example, the profit yielded by taking a short arbitrage position is $-\epsilon$.

The terminal condition is

$$V(\epsilon, I, T) = V(0, I, T) = 0$$

When $\epsilon = \epsilon_{\min}$ and $\epsilon = \epsilon_{\max}$ ($\epsilon_{\max} > C_2$), the boundary conditions are

$$V(\epsilon_{\min}, I, t) = -\epsilon_{\min} - C_2$$

$$V(\epsilon_{\max}, I, t) = 0$$

When $I = I_{\min}$ and $I = I_{\max}$, the boundary conditions are [1]

$$V(\epsilon, I_{\min}, 0) = V(\epsilon, I_{\max}, 0) = \max\{-\epsilon - C_2, 0\}$$

[1] it is not trivial to give the Dirichlet-type conditions at first. We assume that a small change of I doesn't change the option value V much, hence the Neuman-type boundary conditions are

Incorporating with the other two options, we present the complete model in the following succinct form on $(\epsilon, I, t) \in \{[\epsilon_{\min}, \epsilon_{\max}] \times [I_{\min}, I_{\max}] \times [0, T)\}$

$$\min\{-\mathcal{L}H, H - G_{\text{NP}}\} = 0 \tag{3.4}$$

where G_{NP} is the lower bound function for 'no position limits' case.

$$G_{\text{NP}} = \begin{cases} -\epsilon - C_2 & \text{if } H = V \\ \epsilon - C_2 & \text{if } H = U \\ \max(\epsilon + V, -\epsilon + U) - C_1 & \text{if } H = W \end{cases}$$

with the terminal condition,

$$H(\epsilon = 0, I, t = T) = 0$$

with the boundary conditions,

$$H(\epsilon_{\min}, I, t) = \begin{cases} -\epsilon_{\min} - C_2 & \text{if } H = V \\ 0 & \text{if } H = U \\ -\epsilon_{\min} - C_1 & \text{if } H = W \end{cases}$$

and

$$H(\epsilon_{\max}, I, t) = \begin{cases} 0 & \text{if } H = V \\ \epsilon_{\max} - C_2 & \text{if } H = U \\ \epsilon_{\max} - C_1 & \text{if } H = W \end{cases}$$

and

$$H(\epsilon, I_{\min}, t) = \begin{cases} \max\{-\epsilon - C_2, 0\} & \text{if } H = V \\ \max\{\epsilon - C_2, 0\} & \text{if } H = U \\ \max\{\epsilon - C_1 + V, -\epsilon - C_1 + U, 0\} & \text{if } H = W \end{cases}$$

and

$$H(\epsilon, I_{\max}, t) = H(\epsilon, I_{\min}, t)$$

$\frac{\partial V(\epsilon, I_{\min}, 0)}{\partial I} = \frac{\partial V(\epsilon, I_{\max}, 0)}{\partial I} = 0$. We have solved the PDE by imposing the Neuman boundary conditions, but find the numerical values on $I = I_{\min}$ and $I = I_{\max}$ coincide the Dirichlet boundary conditions.

For '*no position limits*' case, the system of PDEs (3.4) are easy to solve because they are not really 'coupled'. We can solve the first two on their own, then using the results solved by first two variational equations to solve the third variational equation. The three options values do not need to be solved simultaneously.

3.1.3 With Position Limits

Under 'with position limits' assumption, closing an arbitrage position not only yields an profit but also gives the right to initiate a new arbitrage position later on. Therefore, we can identify the lower bounds for three options.

$$\begin{cases} V(\epsilon, I, t) \geq \max\{W(\epsilon, I, t) - \epsilon - C_2, 0\} \\ U(\epsilon, I, t) \geq \max\{W(\epsilon, I, t) + \epsilon - C_2, 0\} \\ W(\epsilon, I, t) \geq \max\{\epsilon + V(\epsilon, I, t) - C_1, -\epsilon + U(\epsilon, I, t) - C_1, 0\} \end{cases}$$

we can present the complete model in the following succinct form on $(\epsilon, I, t) \in \{[\epsilon_{\min}, \epsilon_{\max}] \times [I_{\min}, I_{\max}] \times [0, T)\}$

$$\min\{-\mathcal{L}H, H - G_{\text{WP}}\} = 0 \tag{3.5}$$

where G_{WP} is the lower bound function for 'with position limits' case.

$$G_{\text{WP}} = \begin{cases} W - \epsilon - C_2 & \text{if } H = V \\ W + \epsilon - C_2 & \text{if } H = U \\ \max(\epsilon + V, -\epsilon + U) - C_1 & \text{if } H = W \end{cases}$$

with the terminal condition,

$$H(\epsilon = 0, I, t = T) = 0$$

with the boundary conditions,

$$H(\epsilon_{\min}, I, t) = \begin{cases} -2\epsilon_{\min} - C_1 - C_2 & \text{if } H = V \\ 0 & \text{if } H = U \\ -\epsilon_{\min} - C_1 & \text{if } H = W \end{cases}$$

and

$$H(\epsilon_{\max}, I, t) = \begin{cases} 0 & \text{if } H = V \\ 2\epsilon_{\max} - C_1 - C_2 & \text{if } H = U \\ \epsilon_{\max} - C_1 & \text{if } H = W \end{cases}$$

and

$$H(\epsilon, I_{\min}, t) = \begin{cases} \max\{-2\epsilon - C_1 - C_2, 0\} & \text{if } H = V \\ \max\{2\epsilon - C_1 - C_2, 0\} & \text{if } H = U \\ \max\{\epsilon - C_1 + V, -\epsilon - C_1 + U, 0\} & \text{if } H = W \end{cases}$$

and

$$H(\epsilon, I_{\max}, t) = H(\epsilon, I_{\min}, t)$$

For '*with position limits*' case, the system of PDEs (3.5) is nested, the variational inequality of each option involves the value of at least one other options. We need to solve these options simultaneously at each time step. We adopt an iterative method, and stop the iteration when the value of each option changes is within a preset tolerance in two consecutive iterations.

3.2 Numerical Scheme

The most common way to solve two dimensional parabolic PDE is Alternating Direction Implicit (ADI) method. The advantage of the ADI method is that the equations that have to be solved in every iteration have a simpler structure and are thus easier to solve. The idea behind the ADI method is to split the finite difference equations into two, one with the x-derivative taken implicitly and the next with the y-derivative taken implicitly. It is equivalent to solve two one dimensional PDEs, line by line. Unfortunately, ADI is not applicable here. The reason is that we need to check early exercise, namely intrinsic value and option values, at each time step. Therefore we have to solve the PDE layer by layer. Carefulness must be taken when building the nine-diagonal matrix (not necessarily the $\mathbf{M}$-matrix) and handling the boundary conditions.

The solution region is confined as

$$\Omega = \{(\epsilon, I, t)\,|\epsilon_{\min} \le \epsilon \le \epsilon_{\max}, I_{\min} \le I \le I_{\max}, 0 \le t \le T\} \tag{3.6}$$

The grid for the finite difference scheme is defined as followed:

$$\epsilon_i = \epsilon_{\min} + i \cdot \delta\epsilon, \quad i = 0, 1, \cdots, m, \quad \epsilon_0 = \epsilon_{\min},\ \epsilon_m = \epsilon_{\max}$$

$$I_j = I_{\min} + j \cdot \delta I, \quad j = 0, 1, \cdots, n, \quad I_0 = I_{\min},\ I_n = I_{\max}$$

$$t_k = k \cdot \delta t \quad k = 0, 1, \cdots, l, \quad t_0 = 0,\ t_l = T$$

Define the grid function

$$H = \{H_{i,j,k}\,|0 \le i \le m, 0 \le j \le n, 0 \le k \le l, \} \tag{3.7}$$

where

$$H_{i,j,k} = H(\epsilon_i, I_j, t_k) \text{ for } 0 \le i \le m, 0 \le j \le n, 0 \le k \le l,$$

Equation (3.3) can be discretized by a standard two factor finite difference method with variable timeweighting to give[2]

$$\begin{aligned}
H_{i,j,k+1} + (1-\theta)\,[\,&\alpha_{i,j}H_{i+1,j+1,k+1} + \eta_{i,j}H_{i+1,j,k+1} - \alpha_{i,j}H_{i+1,j-1,k+1} \\
&+ \kappa_{i,j}H_{i,j+1,k+1} - \beta_{i,j}H_{i,j,k+1} + \phi_{i,j}H_{i,j-1,k+1} \\
&- \alpha_{i,j}H_{i-1,j+1,k+1} + \zeta_{i,j}H_{i-1,j,k+1} + \alpha_{i,j}H_{i-1,j-1,k+1}] \\
= H_{i,j,k} + \theta\,[\,&-\alpha_{i,j}H_{i+1,j+1,k} - \eta_{i,j}H_{i+1,j,k} + \alpha_{i,j}H_{i+1,j-1,k} \\
&- \kappa_{i,j}H_{i,j+1,k} + \beta_{i,j}H_{i,j,k} - \phi_{i,j}H_{i,j-1,k} \\
&+ \alpha_{i,j}H_{i-1,j+1,k} - \zeta_{i,j}H_{i-1,j,k} - \alpha_{i,j}H_{i-1,j-1,k}]
\end{aligned}$$

[2]The first order coefficients in PDE (3.3) are $-\frac{(c+d\cdot\mathrm{sgn}(\epsilon)I)\epsilon}{(T-t)}$ and aI whose signs are uncertain. Therefore we need to consider upwinding scheme to ensure the monotonicity property.

$\theta = 1$ for fully implicit scheme, $\theta = 0.5$ for Crank-Nicolson scheme.

where

$$\begin{cases} \alpha_{i,j} = \frac{1}{4}\rho\gamma b \frac{\delta t}{\delta\epsilon\delta I} \\ \beta_{i,j} = \gamma^2 \frac{\delta t}{\delta\epsilon^2} + b^2 \frac{\delta t}{\delta I^2} + r\delta t + \frac{(c+d\cdot\mathrm{sgn}(\epsilon_i)I_j)|\epsilon_i|}{T-t_{k+\frac{1}{2}}} \frac{\delta t}{\delta\epsilon} + a\,|I_j|\,\frac{\delta t}{\delta I} \\ \eta_{i,j} = \frac{1}{2}\gamma^2 \frac{\delta t}{\delta\epsilon^2} - \frac{(c+d\cdot\mathrm{sgn}(\epsilon_i)I_j)\epsilon_i}{T-t_{k+\frac{1}{2}}} \frac{\delta t}{\delta\epsilon} \cdot \mathbf{1}_{[\epsilon_i<0]} \\ \kappa_{i,j} = \frac{1}{2}b^2 \frac{\delta t}{\delta I^2} - aI_j \frac{\delta t}{\delta I} \cdot \mathbf{1}_{[I_j<0]} \\ \zeta_{i,j} = \frac{1}{2}\gamma^2 \frac{\delta t}{\delta\epsilon^2} + \frac{(c+d\cdot\mathrm{sgn}(\epsilon_i)I_j)\epsilon_i}{T-t_{k+\frac{1}{2}}} \frac{\delta t}{\delta\epsilon} \cdot \mathbf{1}_{[\epsilon_i\geq 0]} \\ \phi_{i,j} = \frac{1}{2}b^2 \frac{\delta t}{\delta I^2} + aI_j \frac{\delta t}{\delta I} \cdot \mathbf{1}_{[I_j\geq 0]} \end{cases}$$

For notational convenience, it helps to rewrite the above discrete equations in matrix form. Let

$$\mathbf{H}_{k+1} = [H_{1,1,k+1}, H_{1,2,k+1}, \cdots, H_{1,n-1,k+1}, \cdots, H_{m-1,1,k+1}, H_{m-1,2,k+1}, \cdots, H_{m-1,n-1,k+1}]^T$$

and

$$\mathbf{H}_k = [H_{1,1,k}, H_{1,2,k}, \cdots, H_{1,n-1,k}, \cdots, H_{m-1,1,k}, H_{m-1,2,k}, \cdots, H_{m-1,n-1,k}]^T$$

and we obtain a compact matrix form

$$(\mathbf{I} + \theta\mathbf{M})\,\mathbf{H}_k = (\mathbf{I} - (1-\theta)\mathbf{M})\,\mathbf{H}_{k+1} + \mathbf{B} \tag{3.8}$$

where matrix $\mathbf{I}$ is an identical matrix, vector $\mathbf{B}$ handles the boundary conditions, and nine-diagonal matrix $\mathbf{M}$ is

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_1(1) & \mathbf{M}_1(2) & & & \\ \mathbf{M}_2(3) & \mathbf{M}_2(1) & \mathbf{M}_2(2) & & \\ & \ddots & \ddots & \ddots & \\ & & \mathbf{M}_{m-2}(3) & \mathbf{M}_{m-2}(1) & \mathbf{M}_{m-2}(2) \\ & & & \mathbf{M}_{m-1}(3) & \mathbf{M}_{m-1}(1) \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} \mathbf{B}(1) \\ \mathbf{B}(2) \\ \vdots \\ \mathbf{B}(m-2) \\ \mathbf{B}(m-1) \end{bmatrix}$$

It is important to note that $\mathbf{M}$ and $\mathbf{B}$ are block matrices. We have three type sub-matrices and use the number in bracket to denote its type. For example $\mathbf{M}(1)$

is the type 1 matrix. Then we can decompose $\mathbf{M}_i(p)$ $(i = 1, 2, \cdots, m-1)$ and $p = 1, 2, 3$

$$\mathbf{M}_i(1) = \begin{bmatrix} \beta_{i,1} & -\kappa_{i,1} & & & \\ -\phi_{i,2} & \beta_{i,2} & -\kappa_{i,2} & & \\ & \ddots & \ddots & \ddots & \\ & & -\phi_{i,n-2} & \beta_{i,n-2} & -\kappa_{i,n-2} \\ & & & \phi_{i,n-1} & \beta_{i,n-1} \end{bmatrix}$$

$$\mathbf{M}_i(2) = \begin{bmatrix} -\eta_{i,1} & -\alpha_{i,1} & & & \\ \alpha_{i,2} & -\eta_{i,2} & -\alpha_{i,2} & & \\ & \ddots & \ddots & \ddots & \\ & & \alpha_{i,n-2} & -\eta_{i,n-2} & -\alpha_{i,n-2} \\ & & & \alpha_{i,n-1} & -\eta_{i,n-1} \end{bmatrix}$$

$$\mathbf{M}_i(3) = \begin{bmatrix} -\zeta_{i,1} & \alpha_{i,1} & & & \\ -\alpha_{i,2} & -\zeta_{i,2} & \alpha_{i,2} & & \\ & \ddots & \ddots & \ddots & \\ & & -\alpha_{i,n-2} & -\zeta_{i,n-2} & \alpha_{i,n-2} \\ & & & -\alpha_{i,n-1} & -\zeta_{i,n-1} \end{bmatrix}$$

and decompose $\mathbf{B}$ to

$$\mathbf{B}(i) = \theta \mathbf{B}^k(i) + (1-\theta)\mathbf{B}^{k+1}(i)$$

where for $i = 1, 2, \cdots, m-1$ and $q = k, k+1$

$$\mathbf{B}^q(i) = \begin{cases} \begin{bmatrix} \alpha_{i,1}H_{i-1,0,q} + \zeta_{i,1}H_{i-1,1,q} - \alpha_{i,1}H_{i-1,2,q} \\ \alpha_{i,2}H_{i-1,1,q} + \zeta_{i,2}H_{i-1,2,q} - \alpha_{i,2}H_{i-1,3,q} \\ \vdots \\ \alpha_{i,n-2}H_{i-1,n-3,q} + \zeta_{i,n-2}H_{i-1,n-2,q} - \alpha_{i,n-2}H_{i-1,n-1,q} \\ \alpha_{i,n-1}H_{i-1,n-2,q} + \zeta_{i,1}H_{i-1,n-1,q} - \alpha_{i,n-1}H_{i-1,n,q} \end{bmatrix} & i = 1 \\ \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} & i = 2, \cdots, m-2 \\ \begin{bmatrix} -\alpha_{i,1}H_{i+1,0,q} + \zeta_{i,1}H_{i+1,1,q} + \alpha_{i,1}H_{i+1,2,q} \\ -\alpha_{i,2}H_{i+1,1,q} + \zeta_{i,2}H_{i+1,2,q} + \alpha_{i,2}H_{i+1,3,q} \\ \vdots \\ -\alpha_{i,n-2}H_{i+1,n-3,q} + \zeta_{i,n-2}H_{i+1,n-2,q} + \alpha_{i,n-2}H_{i+1,n-1,q} \\ -\alpha_{i,n-1}H_{i+1,n-2,q} + \zeta_{i,n-1}H_{i+1,n-1,q} + \alpha_{i,n-1}H_{i+1,n,q} \end{bmatrix} & i = m-1 \end{cases}$$

The matrix $\mathbf{I} + \theta\mathbf{M}$ is a row diagonally dominant matrix, hence the projected SOR ensures the convergence of the numerical solutions. The overrelaxtion method should take into account the nine-diagonal nature of the matrix $\mathbf{I} + \theta\mathbf{M}$, and it should also be adjusted for early exercise. Let $g_{i,j,k}$, for $i = 1, 2, \cdots, m-1$ and $j = 1, 2, \cdots n-1$, be the intrinsic value when $\epsilon = \epsilon_i$ and $I = I_j$. Therefore,

$$g_{i,j,k} = \begin{cases} \max\{-\epsilon_i - C_2, 0\} & \text{if } H = V \\ \max\{\epsilon_i - C_2, 0\} & \text{if } H = U \\ \max\{\max(\epsilon_i + V_{i,j,k}, -\epsilon_i + U_{i,j,k}) - C_1, 0\} & \text{if } H = W \end{cases}$$

and we denote the right hand of equation (3.8)

$$\mathbf{z}_{k+1} = (\mathbf{I} - (1-\theta)\mathbf{M})\mathbf{H}_{k+1} + \mathbf{B}$$

For each time layer k, let $\mathbf{H}_k^l$ be the lth estimate for $\mathbf{H}_k$, the project SOR method for '*no position limits*' case can then be written as in Algorithm 3.

Algorithm 3 Pseudo Code of Projected SOR Method for No Position Limits, Two Dimensional Problem. Determining Option Values $H_{i,j,k}$ for Interior Node (ϵ_i, I_j, t_k)

Let $\mathbf{H}_k^0 = \mathbf{H}_{k+1}$
for $l = 0, 1, 2 \cdots$ until convergence **do**
  if $i = 1$
    if $j = 1$
    $$H_{i,j,k}^{l+1} = \max\left\{g_{i,j,k}, H_{i,j,k}^l + \tfrac{\omega}{1+\theta(\beta_{i,j}-\phi_{i,j})}\left[z_{i,j,k+1} - (1+\theta(\beta_{i,j} - \phi_{i,j}))H_{i,j,k}^l + \theta\kappa_{i,j}H_{i,j+1,k}^l + \theta\eta_{i,j}H_{i+1,j,k}^l + \theta\alpha_{i,j}H_{i+1,j+1,k}^l\right]\right\}$$
    elseif $j = 2 : n-2$
    $$H_{i,j,k}^{l+1} = \max\left\{\begin{array}{l} g_{i,j,k}, H_{i,j,k}^l + \tfrac{\omega}{1+\theta\beta_{i,j}}\left[z_{i,j,k+1} - (1+\theta\beta_{i,j})H_{i,j,k}^l + \theta\phi_{i,j}H_{i,j-1,k}^{l+1}\right. \\ \left. -\theta\alpha_{i,j}H_{i+1,j-1,k}^l + \theta\kappa_{i,j}H_{i,j+1,k}^l + \theta\eta_{i,j}H_{i+1,j,k}^l + \theta\alpha_{i,j}H_{i+1,j+1,k}^l\right] \end{array}\right\}$$
    elseif $j = n-1$
    $$H_{i,j,k}^{l+1} = \max\left\{g_{i,j,k}, H_{i,j,k}^l + \tfrac{\omega}{1+\theta(\beta_{i,j}-\kappa_{i,j})}\left[z_{i,j,k+1} - (1+\theta(\beta_{i,j} - \kappa_{i,j}))H_{i,j,k}^l + \theta\phi_{i,j}H_{i,j-1,k}^{l+1} - \theta\alpha_{i,j}H_{i+1,j-1,k}^l + \theta\eta_{i,j}H_{i+1,j,k}^l\right]\right\}$$
    end if
  elseif $i = 2 : m-2$
    if $j = 1$
    $$H_{i,j,k}^{l+1} = \max\left\{\begin{array}{l} g_{i,j,k}, H_{i,j,k}^l + \tfrac{\omega}{1+\theta(\beta_{i,j}-\phi_{i,j})}\left[z_{i,j,k+1} - (1+\theta(\beta_{i,j} - \kappa_{i,j}))H_{i,j,k}^l + \theta\zeta_{i,j}H_{i-1,j,k}^{l+1}\right. \\ \left. -\theta\alpha_{i,j}H_{i-1,j+1,k}^{l+1} + \theta\kappa_{i,j}H_{i,j+1,k}^l + \theta\eta_{i,j}H_{i+1,j,k}^l + \theta\alpha_{i,j}H_{i+1,j+1,k}^l\right] \end{array}\right\}$$
    elseif $j = 2 : n-2$
    $$H_{i,j,k}^{l+1} = \max\left\{\begin{array}{l} g_{i,j,k}, H_{i,j,k}^l + \tfrac{\omega}{1+\theta\beta_{i,j}}\left[z_{i,j,k+1} - (1+\theta\beta_{i,j})H_{i,j,k}^l + \theta\alpha_{i,j}H_{i-1,j-1,k}^{l+1} - \theta\alpha_{i,j}H_{i-1,j+1,k}^{l+1}\right. \\ \left. +\theta\kappa_{i,j}H_{i,j+1,k}^l + \theta\zeta_{i,j}H_{i-1,j,k}^{l+1} + \theta\phi_{i,j}H_{i,j-1,k}^{l+1} - \theta\alpha_{i,j}H_{i+1,j-1,k}^l + \theta\eta_{i,j}H_{i+1,j,k}^l + \theta\alpha_{i,j}H_{i+1,j+1,k}^l\right] \end{array}\right\}$$
    elseif $j = n-1$
    $$H_{i,j,k}^{l+1} = \max\left\{\begin{array}{l} g_{i,j,k}, H_{i,j,k}^l + \tfrac{\omega}{1+\theta(\beta_{i,j}-\kappa_{i,j})}\left[z_{i,j,k+1} - (1+\theta(\beta_{i,j} - \kappa_{i,j}))H_{i,j,k}^l + \theta\alpha_{i,j}H_{i-1,j-1,k}^{l+1}\right. \\ \left. +\theta\zeta_{i,j}H_{i-1,j,k}^{l+1} + \theta\phi_{i,j}H_{i,j-1,k}^{l+1} - \theta\alpha_{i,j}H_{i+1,j-1,k}^l + \theta\eta_{i,j}H_{i+1,j,k}^l\right] \end{array}\right\}$$
    end if
  elseif $i = m-1$
    if $j = 1$
    $$H_{i,j,k}^{l+1} = \max\left\{g_{i,j,k}, H_{i,j,k}^l + \tfrac{\omega}{1+\theta(\beta_{i,j}-\phi_{i,j})}\left[z_{i,j,k+1} - (1+\theta(\beta_{i,j} - \phi_{i,j}))H_{i,j,k}^l + \theta\zeta_{i,j}H_{i-1,j,k}^{l+1} - \theta\alpha_{i,j}H_{i-1,j+1,k}^{l+1} + \theta\kappa_{i,j}H_{i,j+1,k}^l\right]\right\}$$
    elseif $j = 2 : n-2$
    $$H_{i,j,k}^{l+1} = \max\left\{\begin{array}{l} g_{i,j,k}, H_{i,j,k}^l + \tfrac{\omega}{1+\theta\beta_{i,j}}\left[z_{i,j,k+1} - (1+\theta\beta_{i,j})H_{i,j,k}^l + \theta\alpha_{i,j}H_{i-1,j-1,k}^{l+1}\right. \\ \left. +\theta\zeta_{i,j}H_{i-1,j,k}^{l+1} - \theta\alpha_{i,j}H_{i-1,j+1,k}^{l+1} + \theta\phi_{i,j}H_{i,j-1,k}^{l+1} + \theta\kappa_{i,j}H_{i,j+1,k}^l\right] \end{array}\right\}$$
    elseif $j = n-1$
    $$H_{i,j,k}^{l+1} = \max\left\{g_{i,j,k}, H_{i,j,k}^l + \tfrac{\omega}{1+\theta(\beta_{i,j}-\kappa_{i,j})}\left[z_{i,j,k+1} - (1+\theta(\beta_{i,j} - \kappa_{i,j}))H_{i,j,k}^l + \theta\zeta_{i,j}H_{i-1,j,k}^{l+1} + \theta\alpha_{i,j}H_{i-1,j-1,k}^{l+1} + \theta\phi_{i,j}H_{i,j-1,k}^{l+1}\right]\right\}$$
    end if
  end if
  if $\left\|\mathbf{H}_k^{l+1} - \mathbf{H}_k^l\right\| < tolerance$ **then**
    Quit the iterations
  end if
end for

In Algorithm 3, we just replace H with V, U and W, solve V and U independently, and use the results to solve W finally. However, for '*with position limits*' case, we need solve V, U and W simultaneously. The intrinsic value for them are[3]

$$\begin{cases} g_{i,j,k}^V = \max\{W_{i,j,k} - \epsilon_i - C_2, 0\} \\ g_{i,j,k}^U = \max\{W_{i,j,k} + \epsilon_i - C_2, 0\} \\ g_{i,j,k}^W = \max\{\max(\epsilon_i + V_{i,j,k}, -\epsilon_i + U_{i,j,k}) - C_1, 0\} \end{cases}$$

[3]I use the superscript to denote the corresponding intrinsic value

For each time layer k, let $\mathbf{V}_k^l$, $\mathbf{U}_k^l$ and $\mathbf{W}_k^l$ be the lth estimate for $\mathbf{V}_k$, $\mathbf{U}_k$ and $\mathbf{W}_k$. We present the projected SOR method for '*with position limits*' case which can then be written as in Algorithm 4.

Algorithm 4 Pseudo Code of Projected SOR Method for With Position Limits, Two Dimensional Problem. Determining Option Values $H_{i,j,k}$ for Interior Node (ϵ_i, I_j, t_k)

Let $\mathbf{V}_k^0 = \mathbf{V}_{k+1}$, $\mathbf{U}_k^0 = \mathbf{U}_{k+1}$ and $\mathbf{W}_k^0 = \mathbf{W}_{k+1}$
for $i = 1 : m-1$, $j = 1 : n-1$
  $g_{i,j,k}^V = \max\{W_{i,j,k+1} - \epsilon_i - C_2, 0\}$
  $g_{i,j,k}^U = \max\{W_{i,j,k+1} + \epsilon_i - C_2, 0\}$
end for
for $l = 0, 1, 2 \cdots$ until convergence **do**
  By Algorithm 3: Calculate $V_{i,j,k}^{l+1}$ and $U_{i,j,k}^{l+1}$
  for $i = 1 : m-1$, $j = 1 : n-1$
    $g_{i,j,k}^W = \max\left\{\max(\epsilon_i + V_{i,j,k}^{l+1}, -\epsilon_i + U_{i,j,k}^{l+1}) - C_1, 0\right\}$
  end for
  By Algorithm 3: Calculate $W_{i,j,k}^{l+1}$
  for $i = 1 : m-1$, $j = 1 : n-1$
    $g_{i,j,k}^V = \max\left\{W_{i,j,k+1}^{l+1} - \epsilon_i - C_2, 0\right\}$
    $g_{i,j,k}^U = \max\left\{W_{i,j,k+1}^{l+1} + \epsilon_i - C_2, 0\right\}$
  end for
  if $\left\|\left(\mathbf{V}_k^{l+1} - \mathbf{V}_k^l\right) + \left(\mathbf{U}_k^{l+1} - \mathbf{U}_k^l\right) + \left(\mathbf{W}_k^{l+1} - \mathbf{W}_k^l\right)\right\| < tolerance$ **then**
    Quit the iterations
  end if
end for

3.3 Experimental Results

In this section, we calculate the options price of $V(\epsilon, I, t), U(\epsilon, I, t)$ and $W(\epsilon, I, t)$ and draw their respective early exercise boundaries.

3.3.1 Data Inputs

All data inputs are as the same as in Table 2.1 except μ and ρ. In one dimensional case, μ is a constant and we set $\mu = 0.03$ for numerical experiment. However, in two dimensional case, μ is a function of I and ϵ, precisely, $\mu = c + d \cdot \text{sgn}(\epsilon)I$. When $|I| = 0$, it reduces to one dimensional problem. Without loss of generosity, we set $c = 0.2$ and $d = \frac{c}{I_{\max+\delta i}}$ to make sure positivity of μ.[4]. We partition $N_\epsilon = N_I = N_t = 50$ in state and time variables, and we choose fully implicit scheme for numerical experiment.

Input Parameter	
Standard Deviation γ	0.6
Riskless Interest Rate r	0.07
Time to Maturity T	1
Cost C_1, C_2	1.2, 0.5
Correlation Coefficient ρ	0
a	$2\delta i$
b	$2\delta i$
c	0.2
d	$\frac{c}{I_{\max+\delta i}}$

Table 3.1: *Model Parameters for Stylized Two Dimensional Problem*

3.3.2 Options Value

In the following, Figure 3.1 displays the option values of V, U, and W at $t = 0$ under 'no position limits' case. Use plane $I = a$, ($I_{\min} \leq a \leq I_{\max}$) to cut the three plots, we collected three types of one-dimensional data curve. Given the same

[4] $I = 0$ means there is a balance in actual order book in market, $\mu = c + d \cdot \text{sgn}(\epsilon)I = c + 0 = 0.2$

point of order imbalance I, the cross-sectional datum of three options display a very similar plot as one dimensional curve, shown in Figure 2.6. Firstly, if we rotate the plot of V to 180 degree, it will look very similar to the plot of U. For option value V, it is decreasing with respect to ϵ. Secondly, the curve of W is symmetrical with respect to the axis $\epsilon = 0$. The plot of option values for 'with position limits'

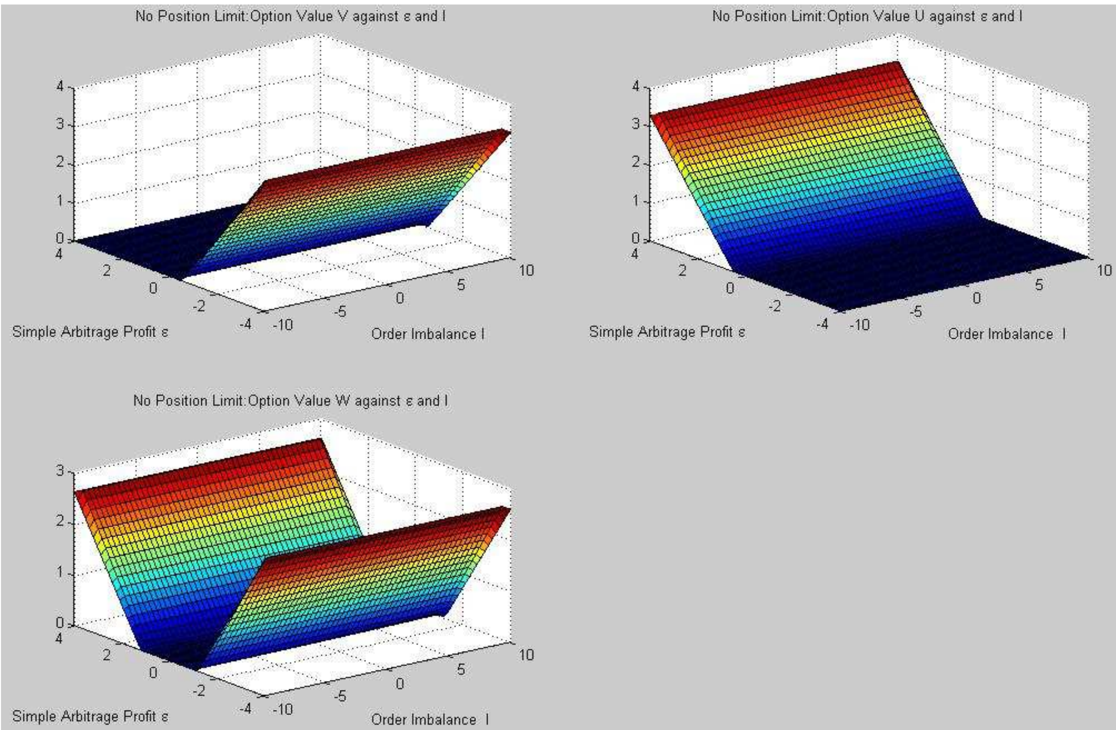

Figure 3.1: *Option Values for V, U, W, No Position Limits, 2D*

case is displayed in Figure 3.2. It is interesting to note that the option V and U are more valuable than their counterparts for 'no position limits' case, while the option W is exactly the same under both cases. We might be fooled by the graphs, hence we conduct a numerical test for option W. We construct a mesh grid, by selecting the points $\epsilon = [-3, -2.5, -2, -1.5, -1, -0.5, 0, 0.5, 1, 1.5, 2, 2.5, 3]$ and $I = [-8, -5, -2, 2, 5, 8]$. W_{NP} and W_{WP} are calculated on these 78 nodes and the norm is

$$\|W_{\text{NP}} - W_{\text{WP}}\| = 9.9895 \times 10^{-7}$$

This shows that the option W are equally valuable regardless to position limits.

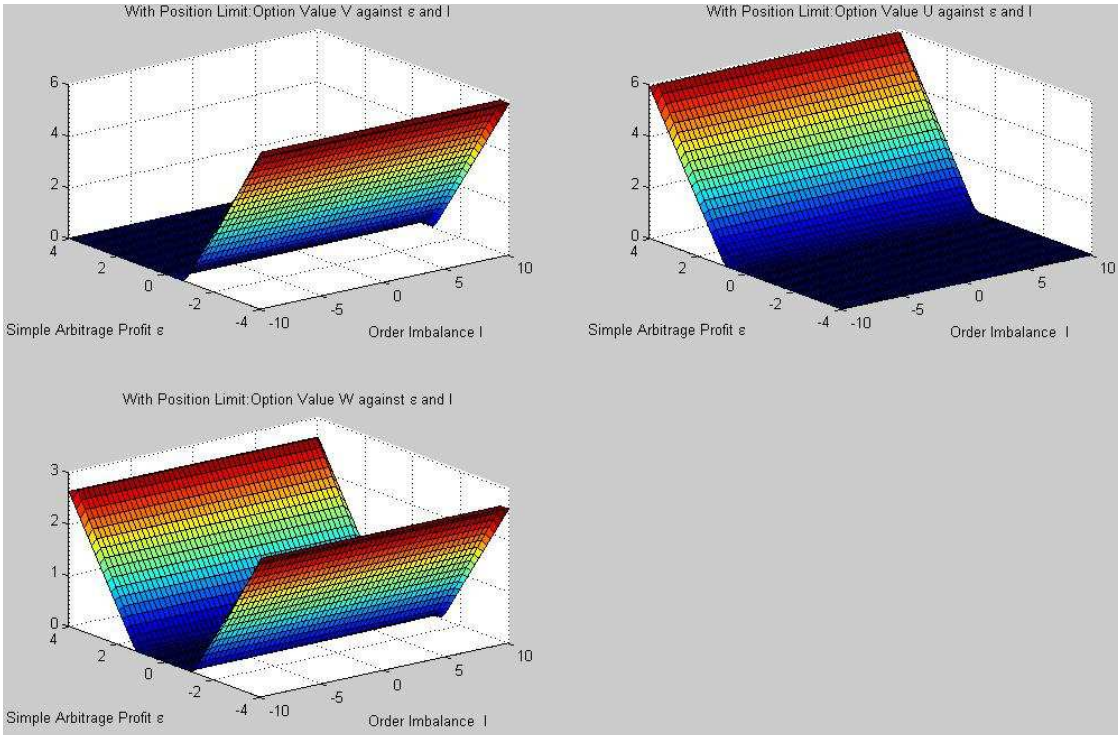

Figure 3.2: *Option Values for V, U, W, With Position Limits, 2D*

Moreover, from theoretical point of view,

1. **Value of V and U are larger for '*with position limits*' case**

   List out the variational equations for both cases.

   For '*no position limits*' case

$$\min\{-\mathcal{L}V, V - (-\epsilon - C_2)\} = 0$$
$$\min\{-\mathcal{L}U, U - (\epsilon - C_2)\} = 0$$

   For '*with position limits*' case

$$\min\{-\mathcal{L}V, V - (W - \epsilon - C_2)\} = 0$$
$$\min\{-\mathcal{L}U, U - (W + \epsilon - C_2)\} = 0$$

Clearly, the options of V and U for 'with position limits' case have an additional non-negative value in the lower bound.

2. **Value of W is exactly the same for both cases**

   In both cases we have

$$\min\{-\mathcal{L}W, W - \max(\epsilon + V - C_1, -\epsilon + U - C_1)\} = 0$$

When early exercise happens for W, W can either takes the value of $\epsilon+V-C_1$ or $-\epsilon+U-C_1$. When W takes $\epsilon+V-C_1$, which means ϵ is positive in large, hence V goes to zero. In another hand, when W takes $-\epsilon+U-C_1$, which means ϵ is negative in large, hence U approaches to zero. So effectively, the variational inequality of W reduced to

$$\min\{-\mathcal{L}W, W - \max(\epsilon - C_1, -\epsilon - C_1)\} = 0$$

The variational equation above indicates that W is independent of values of V and U. Therefore, for both models, W are identical although V and U have different values. Economically speaking, it means the value of the option for investor to initiate an arbitrage position is not affected by existence of position limits.

3.3.3 Early Exercise Boundary

We first consider 'no position limits' case. Figure 3.3 displays the early exercise boundary for option V when the order imbalance changes. For option V, we will hold it when $V > -\epsilon - C_2$ since the option value is more than the payoff obtained when option is immediately exercised. Therefore, the left part is the exercise region and the right part is the holding region. In this figure, we draw six early exercise boundaries given $I = -8, -5, -2, 2, 5, 8$. Although the boundary is discontinuous which might be due to a coarse discretization on $\epsilon - I$ grid, it is approaching to

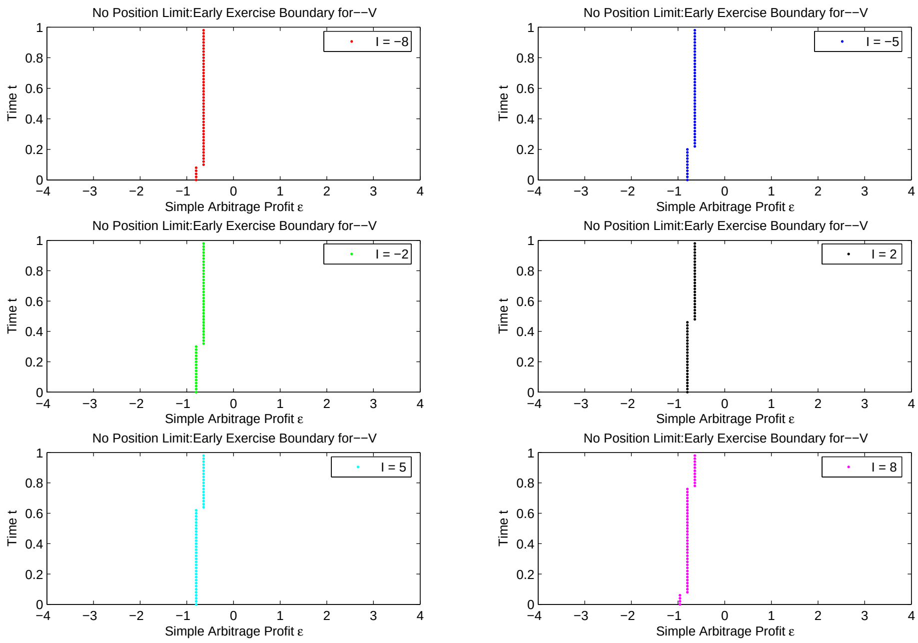

Figure 3.3: *Early Exercise Boundary of Option V, for different values of I*

the left when I is increasing. In another words, the exercise region become smaller and holding region become bigger when I increases. This is intuitively correct. A positive I means more people are buying the futures than selling the futures, which indicates that more people want to lock in the arbitrage profit $-\epsilon$ (from buying the futures) than ϵ (from selling the futures). Hence the option V is more valuable for a larger I than the one for a smaller I, so the holding region should be bigger and the exercise region should be smaller.

Figure 3.4 displays the early exercise boundary for option U when the order imbalance changes. For option U, we will hold it when $U > \epsilon - C_2$ since the option value is more than the payoff obtained when option is immediately exercised. Therefore, the left part is the holding region and the right part is the exercise region. In this figure, we draw six early exercise boundaries given $I = -8, -5, -2, 2, 5, 8$.

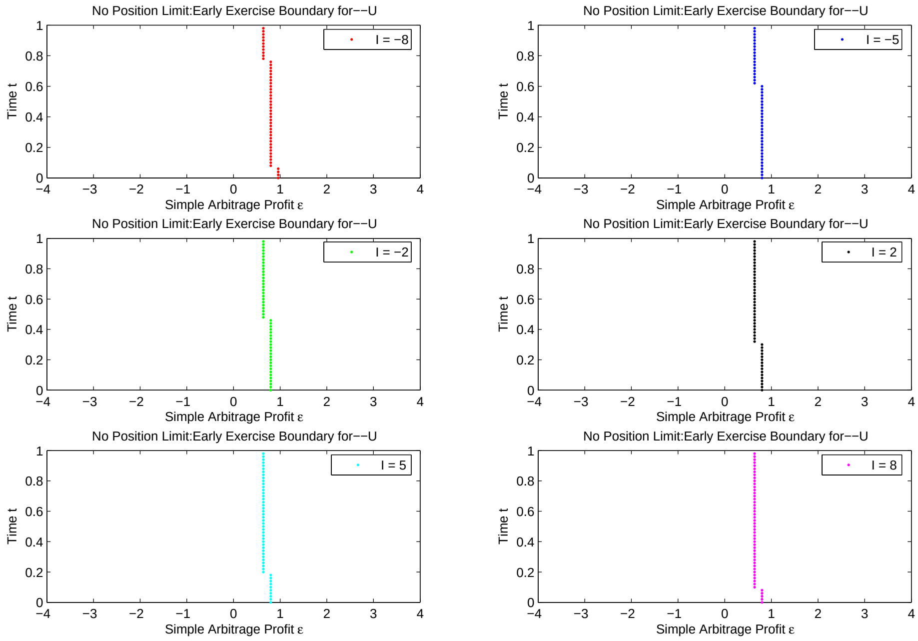

Figure 3.4: *Early Exercise Boundary of Option U, for different values of I*

Although the boundary is discontinuous which might be due to a coarse discretization on $\epsilon - I$ grid, it is also approaching to the left when I is increasing. In another words, the exercise region become smaller and holding region become bigger when I decreases. This is intuitively correct. A negative I means more people are selling the futures than buying the futures, which indicates that more people want to lock in the arbitrage profit ϵ (from selling the futures) than $-\epsilon$ (from buying the futures). Hence the option U is more valuable for a smaller I than the one for a larger I, so the holding region should be bigger and the exercise region should be smaller.

Figure 3.5 displays the early exercise boundary for option U when the order imbalance changes. For option W, we will hold it when $W > \max(\epsilon + V, -\epsilon + U) - C_1$ since the option value is more than the payoff obtained when option is immediately exercised. Therefore, the middle part is the holding region and the left and right

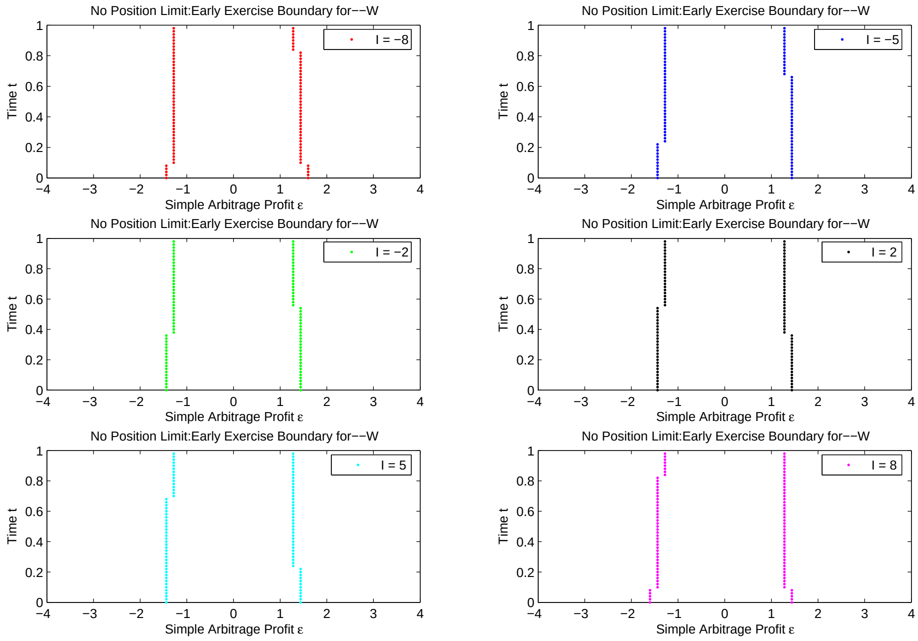

Figure 3.5: *Early Exercise Boundary of Option W, for different values of I*

parts are the exercise regions. In this figure, we draw six early exercise boundaries given $I = -8, -5, -2, 2, 5, 8$. There are two early exercise boundaries for option W, the left one is contributed by option V and the right one is contributed by option U. When I is increasing, both early exercise boundaries move to the left. The reason has been explained above. Intriguingly, we find the early exercise boundary shows a symmetrical pattern for $I = \pm 8$, $I = \pm 5$ and $I = \pm 2$. In another words, the holding region for option W is identical for those values of I which have the same magnitude but opposite sign. Moreover, from the plot, we conjecture that the area of the holding region is unchanged regardless to the value of I.

Similar to the one dimensional problem, the early exercise boundaries for three options are exactly the same, independent of position limits. The next three figures depict the identity.

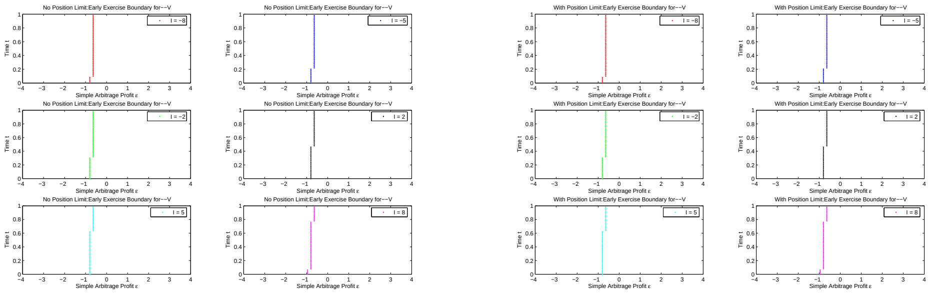

Figure 3.6: *For Both Cases: The Early Exercise Boundaries of Option V*

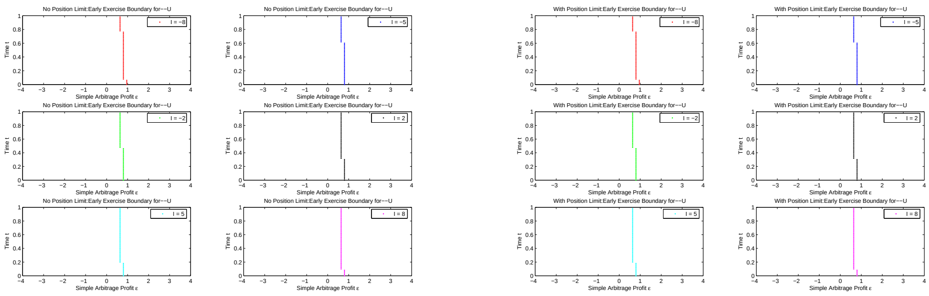

Figure 3.7: *For Both Cases: The Early Exercise Boundaries of Option U*

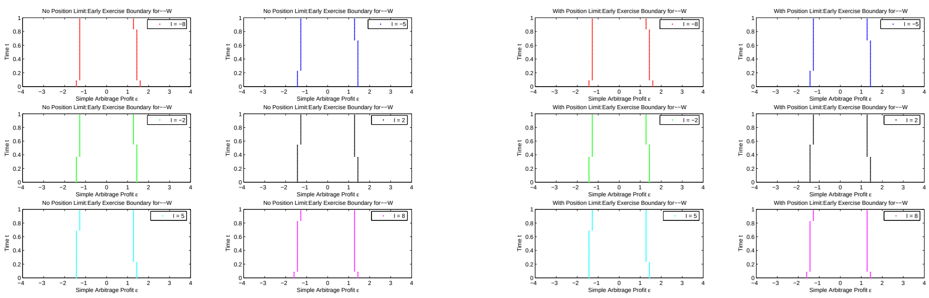

Figure 3.8: *For Both Cases: The Early Exercise Boundaries of Option W*

Chapter 4

Conclusion

In this thesis, we mainly focus on pricing options whose payoff is based on simple arbitrage profit in stock index futures and plotting their early exercise boundaries. We consider both one dimensional and two dimensional problems, for each we subdivide as 'no position limits' case and 'with position limits' case.

In one dimensional problem, we use Brownian Bridge process to model simple arbitrage profit. A one dimensional PDE for the options is derived. In two dimensional problem, we add one mean-reverting stochastic differential equation to model order imbalance. A two dimensional PDE for the options is derived. We also take into account of transaction costs and position limits and form complete models.

We use fully implicit and Crank-Nicolson scheme to solve the variational inequality numerically. To handle American option type, we adopt projected SOR method. Numerical Results of the early exercise boundaries and option values are given and analyzed. These early exercise boundaries give us the optimal arbitrage strategy. We discuss various parameter effects on option values and early exercise boundary, for one dimensional problem, while we also examine the order imbalance impacts on early exercise boundary, for two dimensional problem. We also compare the numerical results between the 'no position limits' and 'with position limits' models, and find the optimal trading strategy is exactly the same for both cases.

Two possible future works can be extended as follows.

- We can conduct an empirical study by gathering numerous financial data, such as index futures price, dividend and order positions from Bloomberg.
- Although the early exercise boundaries are derived and analyzed numerically, we could mathematically analyze the properties of the early exercise boundaries.

Bibliography

[1] Cornell, B., and French, K. Taxes and the Pricing of Stock Index Futures. *Journal of Finance*, 1983, vol. 38, issue 3, pages 675-694

[2] Figlewski, S. 1985. Hedging with Stock Index Futures: Theory and Application in a New Market. *Journal of Futures Markets* 5:183-200

[3] Brennan, M., and Schwartz, E. Arbitrage in Stock Index Futures. *The Journal of Business, Vol. 63, No. 1, Part 2: A Conference in Honor of Merton H. Miller's Contributions to Finance and Economics* (Jan., 1990), pp. S7-S31

[4] Fund, J.K.W. (2007). Order imbalance and the pricing of index futures. *Journal of Futures Markets* 27, 697-717

[5] Joseph K.W. Fung, Philip L.H. Yu. (2008) Order imbalance and the dynamics of index and futures prices. *Journal of Futures Markets* 27:12, 1129-1157

[6] Chen Huan. (2007) Optimal Arbitrage Strategy in Stock Index Futures *Final Year Project*

[7] M. Dai, Y.K. Kwok, and Y.F. Zhong (2009) Optimal arbitrage strategies on stock under index futures and under portion limits. *working paper*

[8] Michael Monoyios, Lucio Sarno. (2002) Mean reversion in stock index futures markets: A nonlinear analysis. *Journal of Futures Markets* 22:4, 285-314

[9] Fima C Klebaner. (2006) Introduction to stochastic calculus with applications. Second edition. *Imperial College Press*

Appendix A

Appendix

A.1 Analytical Formula of Brownian Bridge

The SDE is

$$d\epsilon = -\frac{\mu\epsilon}{T-t}dt + \gamma dW$$

which is a particular type of general linear SDE

$$d\epsilon(t) = (\alpha(t) + \beta(t)\epsilon(t))dt + (\gamma(t) + \delta(t)\epsilon(t))dW_t$$

with parameters

$$\begin{cases} \alpha(t) = 0 \\ \beta(t) = -\frac{\mu}{T-t} \\ \gamma(t) = \gamma \\ \delta(t) = 0 \end{cases}$$

In [9], the analytical formula of this general linear type SDE is given

$$\epsilon(t) = U(t)\left(\epsilon(0) + \int_0^t \frac{\alpha(s) - \delta(s)\gamma(s)}{U(s)}ds + \int_0^t \frac{\gamma(s)}{U(s)}dW_s\right) \tag{A.1}$$

where

$$U(t) = e^{\int_0^t \left(\beta(s) - \frac{\delta(s)^2}{2}\right)ds + \int_0^t \delta(s)dW_s} \tag{A.2}$$

Plug the four parameters to the A.1 and A.2, we obtain that

$$\epsilon(t) = \left(1 - \frac{t}{T}\right)^{\mu} \epsilon(0) + \gamma(T-t)^{\mu} \int_0^t \frac{1}{(T-s)^{\mu}} dW_s \qquad 0 \leq t < T \tag{A.3}$$

Since the function under the Itô integral is deterministic, and for any $t < T$, $\int_0^t \frac{1}{(T-s)^{2\mu}} ds < \infty$ when $\mu \neq 0$, the process $\int_0^t \frac{1}{(T-s)^{\mu}} dW_s$ is a martingale, moreover, it is Gaussian. Thus $\epsilon(t)$, on $[0, T)$ is a Gaussian process with initial value $\epsilon(0)$. The value at T, which is $\epsilon(T) = 0$, is determined by continuity by showing that $\lim_{t \uparrow T} \int_0^t \frac{1}{(T-s)^{\mu}} dW_s = 0$ almost surely.