The Evaluation of Gas Sales Agreements

Wenfeng DONG

PhD

University of York
Mathematics

June 2018

Abstract

A gas sales agreement, also called a gas swing contract, is an agreement between a supplier and a purchaser for the delivery of variable daily quantities of gas, between specified minimum and maximum daily limits, over a certain number of years at a strike price. The main constraint of such an agreement is that there is a minimum volume of gas for which the buyer will be charged at the end of the year, regardless of the actual quantity of gas taken. For multiple year contracts, there are also features called the make-up and carry-forward banks which add another level of complexity to the analysis. We propose a framework for pricing such multiple year contracts where both the gas price and strike price are stochastic processes. With the help of a two-dimensional trinomial tree, we are able to price such swing contracts with both make-up and carry-forward banks, and find the optimal daily decisions and the optimal yearly usage of the make-up and carry-forward banks. We also provide a detailed analysis of the different features that these contracts possess. Furthermore, another feature, called the indexation principle, is popular in real markets, under which the strike price is called the index. In each month, the value of the index is determined by the weighted average price of some energy products in the previous month. We design a lattice-based algorithm to price such swing contracts and find optimal daily decisions by using graphics processing units. Since the least-squares Monte Carlo method is well-known to handle sophisticated models, such as multi-factor models, models with regime-switching, or models with jumps, we build this method for the pricing of gas sales agreements and analyze the performance of it, especially the impacts of explanatory variables. With the help of concrete numerical examples, various features of such contracts with indexation are demonstrated.

Acknowledgements

This Ph.D has been an amazing experience for me and this thesis would not have been possible without the support I have received from numerous people.

First of all, I would like to express heartily gratitude to my supervisor, Dr Boda Kang. It has been a great honour to be his first Ph.D student. I sincerely appreciate his continuous support, encouragement, inspiration and guidance throughout my Ph.D. I also appreciate his valuable advice, not only on this thesis, but also on my career. I enjoyed every discussion we have had and the ability for independent, creative and critical thinking that he has embedded in me will be the treasure of my life.

Besides my supervisor, I am especially grateful for my co-supervisor, Professor Zdzislaw Brzezniak, who has offered me this wonderful opportunity to undertake a Ph.D. At the beginning of my Ph.D, I could not imagine how I would manage to get through the hardship but his continual encouragement helped me through these difficult periods. I also want to thank Dr Alet Roux for being the chair of my thesis advisory panel and giving me lots of helpful advice on my research. In addition, thank the Department of Mathematics for the financial support so that I can attend a number of conferences. Furthermore, I would like to thank my examiners, Dr Alet Roux and Professor Tiziano Vargiolu, for their valuable comments on an earlier version and suggestions of changes that helped to improve this thesis.

I would like to thank all my friends and colleagues, in particular, Mr Wenbo Zhou, Miss Paula Avello Fernandez and Mr Zhikang Xu for all their encouragement, interesting discussions and chats.

With love and gratitude, I give thanks to my father, Mr Maoqiang Dong, and mother, Mrs Li'e Jing, for their endless love. Thanks to my little brother, Mr Youbin Dong, for being such an adorable kid while I am away from home.

My special thanks go to my lovely wife, Dr Jiaxi Wang. I thank her for understanding why I have not been able to be with her for most of the time of our marriage while pursuing the Ph.D degree. I thank her for all her love. Without her encouragement and extreme patience this thesis would have never been finished.

Declaration of Authorship

Chapter 3 is based on my paper: "Analysis of a multiple year gas sales agreement with make-up, carry-forward and indexation", joint with Dr Boda Kang.

I, Wenfeng DONG, declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, university. All sources are acknowledged as references.

Contents

List of Figures

List of Tables

Chapter 1

Introduction and literature review

1.1 Introduction

A gas sales agreement (GSA), also termed a gas swing option, or a take-or-pay option, is a contract for the purchase and sale of gas in the natural gas industry. GSAs are designed in such a way as to offer flexibility concerning both the volume and the timing of the delivery of gas. They allow the gas buyer to purchase a quantity of gas from the gas provider periodically in the future at a contractual price. At each delivery date agreed in the contract, the buyer has the obligation to buy a pre-determined base load of gas from the gas provider. Then the buyer has the right to require the delivery of more gas as long as the total gas delivered on this date does not exceed a pre-fixed limit. That is, the buyer has the right to choose both the timing and the quantity of the delivery while the security of the gas provider is protected by the existence of the base load. Besides the limits on each delivery, some global constraints also apply to the total volume of gas delivered through the contract. The global constraints usually require the total volume of gas purchased under the GSA to lie between a pre-determined lower limit (termed the minimum bill) and a pre-determined upper limit (termed the annual contract quantity).

Thanks to the deregulation of the energy market, both the price and the contract details of the GSA are now negotiated between the gas seller and the gas buyer. This makes the understanding and valuation of the optionality built in the GSA more important than before. In addition, unlike a typical

European option or American option, the gas buyer in a GSA contract has multiple opportunities to demand the delivery of gas. Due to the existence of the global constraints, on each agreed contract date, the gas buyer should make a decision as to the volume of gas to be delivered considering not only the current profits but also the impact of this decision on future dates. That is, the understanding of the GSA is not only about the contract price, but also about the impact of the contract features on the volume of gas which should be taken to maximize the profit of the buyers.

The GSA usually offers the gas buyer daily (as well as weekly or monthly) opportunities to purchase gas from the gas seller. That is, regardless of the current market price of the natural gas, on each day, the gas buyer has to buy a base load of gas (called the daily minimum) at a contractual price while making a decision as to whether he/she should ask for a extra purchase of gas under the GSA. Due to the existence of this base load of gas, the GSA can be decomposed into two parts: a swap part and a swing part. The swap part makes sure the gas seller can sell a base load of gas every day. Since daily exercise opportunities are offered, there is no early exercise feature involved in the swap part. The swing part, also called the normalized GSA, gives the buyer the right to purchase extra quantities of gas up to a daily limit (called the daily maximum). This element of the agreement introduces early exercise features and makes the swing part the main challenge in the evaluation of GSAs.

Based on whether the global constraints can be violated or not, GSAs can be divided into two types: GSAs with penalties and GSAs with firm constraints. GSAs with penalties allow the gas buyer to violate the global constraints. That is, the total volume of gas purchased under the GSA can be below the minimum bill or above the annual contract quantity. Once the global constraints are violated, however, the gas buyer has to pay some penalties to the gas seller at the end of each year. The value of the penalties to be paid is also negotiated between the gas buyer and the gas seller, but is often linked to the extra or shortfall quantity of gas purchased. In contrast, GSAs with firm constraints require the total volume of gas purchased to lie strictly between the minimum bill and the annual contract quantity. Compared with GSAs with firm constraints, GSAs with penalties involve extra cashflow generated by the penalties and hence introduce extra challenge. In this thesis, we

mainly focus on GSAs with penalties.

In addition to the daily minimum constraint, the daily maximum constraint and the global constraints, there are other popular features in GSA contracts: the make-up bank, the carry-forward bank and the indexation principle.

1.1.1 The make-up bank and the carry-forward bank

GSAs usually last multiple years. There are two common and important features in multi-year GSA contracts: the make-up bank and the carry-forward bank. In a year where the GSA contract is out-of-the-money[1] on most days, due to the existence of the minimum bill, the buyer is at risk of facing great losses under the contract. These losses can be from the delivery of gas on those out-of-the-money days, the year-end penalties, or both. In this case, the buyer needs ways to reduce losses, and the so-called make-up bank is introduced. With the make-up bank, the buyer can take some quantities of gas which are less than the minimum bill and pay penalties in an out-of-the-money year, and the shortfall between the actual gas taken and the minimum bill is then added to the make-up bank. In later years, where the gas taken is greater than some reference level[2], the extra gas purchased can be withdrawn from the make-up bank and a refund paid. Whereas the make-up bank encourages the buyer to pay penalties, in the hope that these will be refunded later, the so-called carry-forward bank gives the buyer the right to reduce the minimum bill and hence reduce possible penalties. With the carry-forward bank, if the buyer anticipates that the contract will be out-of-the-money in future years, he/she can take more gas in the current year. If the gas taken is greater than some reference level[3], the excess gas can be added to the carry-forward bank. In later years, where the buyer is under pressure to pay penalties, the gas in the carry-forward bank can be withdrawn to reduce the minimum bill in those years. Both the make-up bank and the

[1] On each day, we say the GSA contract is out-of-the-money if the gas price is below the strike price, and the GSA contract is in-the-money if the gas price is above the strike price.

[2] In this thesis, we consider the make-up and carry-forward banks introduced in Breslin et al. (2008). This reference level is the sum of the minimum bill and the volume of gas available in the carry-forward bank.

[3] This reference level is called the carry-forward base. However, if the buyer is withdrawing gas from the make-up bank and adding gas to the carry-forward bank at the same time, other constraints may apply, see Section 3.3.2.

carry-forward bank offer the buyer opportunities to apply flexible strategies in order to minimize losses, hence the way how they affect the contract values and the trading strategies should be investigated carefully. In addition, the make-up bank and the carry-forward bank introduce extra optionality and two more dimensions, which makes the evaluation of the GSA more challenging.

1.1.2 The indexation principle

Besides the make-up bank and the carry-forward bank, there is one more feature, the indexation, which introduces a further difficulty in the evaluation of the GSA. When the contractual price (the strike price) in a GSA is a constant, the valuation of the GSA is a classical dynamic programming problem. In real contracts, however, the contractual price is set based on the indexation principle, under which the contractual price is called the index. In each month, the value of the index is determined by the weighted average price of some energy substitutions (e.g. crude oil) in the previous month (see Asche, Osmundsen and Tveterås (2002) for details). Under the indexation principle, the value of the GSA contract not only depends on the current state, but also depends on the past values of some other energy products. This feature links the valuation of the GSAs to the moving average problem. So far, however, no effective method has been derived in the literature to value GSAs embedded with moving average features.

1.2 Existing methods

Since the GSA is widely used in the natural gas industry, it has received extensive treatment in the literature, but this work has only focused on the classic GSA contracts, containing neither the make-up and carry-forward banks nor the indexation principle. Furthermore, in most of the existing literature, the contractual price is assumed to be a constant and the GSA contract usually lasts only one year. This section, therefore, reviews these existing methods to evaluate the classic GSA.

1.2.1 Lattice-based methods

The first treatment of gas swing valuations appeared in Thompson (1995), where the author applies a binomial tree method extended from Hull and White (1993) to take-or-pay gas contracts. The spot price in this paper is assumed to be a simple geometric Brownian motion. Lari-Lavassani, Simchi and Ware (2001) develop the idea of the binomial tree by considering a two-factor mean-reverting process. Clewlow, Strickland and Kaminski (2001), meanwhile, provide a discussion of the optimal exercise decisions with the help of a recombining trinomial tree. In Jaillet, Ronn and Tompaidis (2004), the authors proposed a method which values the swing options based on the dynamic programming on a trinomial tree. This method starts from the end point of the contract and works backwards in time by considering a mean-reverting spot price model. In Barrera-Esteve et al. (2006), the authors propose and summarize several methods to evaluate GSAs with penalties using both simulations and dynamic programming techniques. A so-called optimal quantization tree method is also built in Bardou, Bouthemy and Pagès (2009) and Bardou, Bouthemy and Pagès (2010) for mean-reverting spot prices. This method is said to be efficient since the quantization tree only needs to be constructed once if the underlying price model does not change. In Wahab and Lee (2011), the authors assume that the spot price follows a regime switching model, where the volatility can switch between different values based on the state of a hidden Markov chain and thence evaluate the swing option on a pentanomial tree. In Benth, Lempa and Nilssen (2012), the authors study the swing option on electricity markets and show that its value is the solution of a Hamilton–Jacobi–Bellman (HJB) equation. Following the theory built in Benth, Lempa and Nilssen (2012), Edoli (2013) models the gas swing option and solves the HJB equation by using finite difference methods.

1.2.2 Simulation-based methods

The least-squares Monte Carlo simulation method (Longstaff and Schwartz (2001)) is well known in the valuation of American options, and has been extended by many authors to accommodate the evaluation of swing options. In Dörr (2003), the author introduces the least-squares Monte Carlo on the evaluation of swing options. In Thanawalla (2006), the author proposes the

use of non-parametric regression using splines, and evaluates the GSAs with penalties. Meinshausen and Hambly (2004) build the so-called dual-pricing approach to get the upper bound of the swing options when a take-or-pay provision is not included. The dual-pricing approach has been further developed in Aleksandrov and Hambly (2010) and Bender (2011) to price the swing options which allow multiple exercise rights on a single date. Another simulation-based approach can be found in Ibáñez (2004), where the authors apply the technique in Ibáñez and Zapatero (2004) by finding optimal exercise prices before pricing swing options. In a recent work, Hanfeld and Schlüter (2017) analyze the least-squares Monte Carlo method by comparing it with a so-called myopic approach which simply exercises rights if the current payoff is positive.

1.3 Motivation

GSAs embedded with two features have been less highlighted in the literature: multiple year GSAs with make-up and carry-forward banks, and GSAs with indexation.

1.3.1 Multiple year GSAs with make-up and carry-forward banks

We have introduced the basic functionalities of the make-up bank and the carry-forward bank in Section 1.1.1. The multiple year GSAs with these banks have only received a few treatments in the literature. In Breslin et al. (2008), the authors explained the basic features of both the make-up bank and the carry-forward bank. In Holden, Løland and Lindqvist (2011), an algorithm is proposed using the least-squares Monte Carlo method to evaluate GSAs with a carry-forward bank, although the authors only price the contract values and do not extract optimal decisions. The evaluation of GSAs is not only about the contract value, however, but also about finding the optimal daily decisions based on various gas prices and index prices. The first formally quantitative treatment of GSAs containing the make-up provision appears in Edoli et al. (2013). The authors modelled one type of make-up clause and performed a sensitivity analysis of the contract value with respect to various parameters. The GSA in Edoli et al. (2013) uses hard constraints. That is, their GSAs do not have penalties and the volume taken in each year must be

within a pre-specified range. Also, the volume of gas in the make-up bank must be zero at the end of the contract. There are GSA contracts in the market, however, that offer more flexibility (see contracts in Subsection 3.3.2); furthermore, in Edoli et al. (2013), the GSAs contain either the make-up clause or the carry-forward clause, but not both. The second treatment of make-up and carry-forward banks appears in Chiarella, Clewlow and Kang (2016), where the authors evaluate GSAs with both make-up and carry-forward banks in a regime-switching forward price curve model. This paper, however, only evaluates GSAs with constant strike prices. In addition, while the optimal daily decisions and optimal usage of both make-up and carry-forward banks are given in Chiarella, Clewlow and Kang (2016), there is no detailed analysis of how the contract features (daily constraints, global constraints, etc.) affect the values and decisions. Furthermore, in both Chiarella, Clewlow and Kang (2016) and Edoli et al. (2013), the effect of the make-up and carry-forward banks on the optimal decisions is not given. In Chapter 3, therefore, we evaluate multiple year GSAs with both make-up and carry-forward banks, as in Breslin et al. (2008), but with stochastic strike prices, while also providing a detailed analysis of how the make-up bank, the carry-forward bank, the stochastic strike prices and the different parameter settings affect both values and decisions.

1.3.2 GSAs with indexation

As introduced in Section 1.1.2, indexation links the evaluation of GSAs to the moving average problem. In the literature, very few articles discuss the valuation of options embedded with both the moving average feature and the early exercise feature. Among those that do, a common approach is the least-squares Monte Carlo simulation, as in Broadie and Cao (2008). The idea is simply to find the continuation value by performing regression on polynomials of the current realized values of the underlying price (gas price) and the strike price (index). Based on the indexation principle, however, the index is determined by the average of past oil prices, hence these oil prices should also have their impact on the contract value. In Grau (2008) and Dirnstorfer, Grau and Zagst (2013), the authors suggest that, in the least-squares Monte

Carlo simulation, the regression should be performed not only on the underlying price and the strike price, but also on the all past values used to compute the strike price. This action leads to a very high dimensional problem, especially in the evaluation of the GSAs, since the computation of the index needs 30 past oil values (we assume one month contains 30 days). Although both Grau (2008) and Dirnstorfer, Grau and Zagst (2013) propose a technique based on the sparse grid to reduce the number of basis functions used in the least-squares Monte Carlo simulation, they can only solve a problem up to ten dimensions and the computing time is very long. In Bernhart, Tankov and Warin (2011), the authors develop a method which uses a finite-dimensional approximation of the infinite-dimensional dynamics of the moving average process based on weighted Laguerre polynomial expansion. They use several so-called Laguerre processes to approximate the moving average process, and they are able to prove that the accuracy of this approximation is better if more Laguerre processes are used. The regression in the least-squares Monte Carlo simulation is performed on all underlying prices and all realized values of these Laguerre processes. This new methodology, however, can be not only very time-consuming but also memory-consuming, especially when the memory is quite limited. In addition, as reported in Warin (2012), there is no big improvement of the option value compared with the usual regression method in Broadie and Cao (2008).

There are also works which valuing options with moving average features using the lattice-based method. In Kao and Lyuu (2003), the authors price moving average lookback options with the help of a binomial tree. The binomial tree is also used in Dai, Li and Zhang (2010) to price moving average barrier options, while Xu, Hong and Qin (2013) develop the so-called willow tree method, which was first introduced in Curran (2001), to price moving average American options.

There are even fewer studies that discuss the evaluation of GSAs with indexation. In Bernhart (2011), the author uses the methodology built in Bernhart, Tankov and Warin (2011) to price a GSA contract with only one exercise opportunity, which is an American option where the strike price is computed based on the indexation. In the least-squares Monte Carlo simulations, however, since all simulated paths of all underlying processes and Laguerre processes have to be stored, one would have to store the path information on the

hard drive due to the lack of memory on RAM. This would slow the computation speed significantly due to the need for file writing/reading operations. In an example in Bernhart (2011), it took 20 hours to evaluate such a GSA with only one exercise opportunity (normally, GSAs have daily exercise opportunities, that is, 365 early exercise opportunities for a one-year contract). Edoli (2013) implements the GSA evaluation by assuming that the moving average process is Markovian and implemented it by using the finite difference method. To the author's knowledge, no efficient method has been derived without making those assumptions. In Chapter 4, we propose a new numerical method based on a two-dimensional trinomial tree to fill the gap. Since the least-squares Monte Carlo simulation is well-known to handle sophisticated models, such as multi-factor models, models with regime-switching, or models with jumps, we build this method for the pricing of gas sales agreements with indexation and analyze its performance in Chapter 5.

1.4 Structure of the thesis

Before proceeding to the discussion of the make-up bank, the carry-forward bank and the indexation, we give an introduction to the classic GSA which contains a number of common features in Chapter 2. In addition, since the two-dimensional trinomial tree is intensively used in this thesis to evaluate the GSA, we also provide a description on how to construct such a two-dimensional tree in this chapter.

The rest of this thesis can be divided into two parts. The first part includes Chapter 3 which investigates the evaluation of the multiple year GSAs with both make-up and carry-forward banks. With the help of a two-dimensional trinomial tree, we are able to price these complex GSA contracts and the find optimal daily decisions and optimal yearly usage of both the make-up bank and the carry-forward bank. With the help of a number of numerical examples, we also provide a detailed analysis, not only of the different features these contracts have, but also how different model parameters will affect both the optimal value and the optimal decisions. The main contents of Chapter 3 have been published in Dong and Kang (2018).

The second part of this thesis, which includes Chapter 4, Chapter 5 and Chapter 6, focuses on the GSAs with indexation. In Chapter 4, we propose an

algorithm based on a two-dimensional trinomial tree that can be used to evaluate GSAs with indexation. In addition, we also formulate the framework of GSAs with indexation in continuous time and prove that the algorithm we proposed has a first-order consistency to a Hamilton–Jacobi–Bellman equation. In Chapter 5, we generalize and propose several algorithms based on the Monte Carlo simulation to evaluate GSAs with indexation. We also investigate the performance of the least-squares Monte Carlo simulations for the evaluation of the GSA contracts through numerous numerical examples. In Chapter 6, we first introduce the general-purpose computing on graphics processing units and provide a description on how to implement the algorithms in Chapter 4 and Chapter 5 on graphics processing units. Then we investigate the performance of the algorithms proposed in the second part of this thesis through numerical examples.

Finally, in Chapter 7, we draw conclusions and suggest future work.

Chapter 2

Gas sales agreements foundations

In this chapter, we give an introduction to the classic gas sales agreement in Section 2.1. The classic GSA is the GSA embedded with daily constraints and global constraints. The daily constraints require the delivery of daily quantities of gas to be between pre-determined minimum and maximum daily limits. The global constraints require the total volume of gas delivered under the GSA to lie within a pre-determined interval. These GSAs can be divided into two types: GSAs with firm constraints and GSAs with penalties. We introduce these two types of GSAs along with their properties in Section 2.1.1 and 2.1.2, respectively. In addition, since the two-dimensional trinomial tree is intensively used in this thesis, in Section 2.2 we provide a description on how we can construct such a tree for mean-reverting processes.

2.1 Classic gas sales agreements

A GSA between a gas provider and a gas buyer can have many specific features to satisfy the needs of both sides in the agreement. Next, we present an example of a one-year GSA that contains a number of common features.

- The contract lasts $T = 1$ year and there are N time periods, N is a positive integer.
- Let the length of the time interval be $\Delta t = T/N$, then the contract duration is equally spaced into N periods. Let $t_n = n \cdot \Delta t$, $n = 0, \ldots, N$,

then the n-th time interval is given by $[t_{n-1}, t_n]$. Hence we have:

$$0 = t_0 < t_1 < t_2 < \ldots < t_{N-1} < t_N = T.$$

We assume that, at each t_n, $n = 1, \ldots, N$, there is a base load of gas delivered from the gas seller to the gas buyer. In addition, at each t_n, $n = 1, \ldots, N$, the buyer has exactly one opportunity to ask for a delivery of more gas from the gas seller, which amounts to N opportunities for a whole year. We call these opportunities the exercise rights. Usually, GSAs can be exercised monthly, weekly or daily.

- Denote the risk-free rate by r. Let S_n and I_n be the gas price and the strike price (the contractual price) at time t_n, respectively. At this point, we assume that the gas price $(S_n)_{0 \leq n \leq N}$ is a Markovian process defined on a probability space $(\Omega, \mathcal{F}, \mathbb{Q})$ and the strike price I_n is a fixed constant for all $n = 0, 1, \ldots, N$. S_n can be the observation of a continuous-time process at time t_n. If we denote $F^S(t, T)$ as the forward price of gas at time $t \in [0, T]$, then, ideally, the gas price S_n should be given by $S_n = F^S(t_n, t_n)$. Since $F^S(t_n, t_n)$ is not a tradeable instrument, however, in practice one has to consider the day-ahead forward price. That is, the gas price is given by $S_n = e^{-r \cdot \Delta t} F^S(t_n, t_{n+1})$ (see Bardou, Bouthemy and Pagès (2009) and Hinz (2006)).

- Let q_{t_n} be the amount of gas taken at time t_n, $n = 1, \ldots, N$, which is constrained by the minimum daily quantity $q_{\min}$ and the maximum daily quantity $q_{\max}$, that is

$$q_{\min} \leq q_{t_n} \leq q_{\max}, \tag{2.1}$$

where $q_{\min}$ and $q_{\max}$ are constants specified in the contract. In this thesis, q_{t_n} is called the exercise decision, and $q_{\min}$ and $q_{\max}$ are called the daily minimum and the daily maximum, respectively. Indeed, $q_{\min}$ is the base load of gas that has to be delivered at each t_n, $n = 1, \ldots, N$.

- We call the sequence of exercise decisions $(q_{t_1}, q_{t_2}, \ldots, q_{t_N})$ the consumption policy, and denote it by

$$\mathbf{q} = \{q_{t_n}\}_{n=1,2,\ldots,N}.$$

Also we write

$$\sum_{n=n_1}^{n_2} \mathbf{q} = \sum_{n=n_1}^{n_2} q_{t_n}, \text{ and } \sum \mathbf{q} = \sum_{n=1}^{N} q_{t_n},$$

where q_{t_n} is the exercise decision in the consumption policy $\mathbf{q}$ on the n-th period.

- Let Q_{t_n} be the cumulative amount of gas taken before time t_n. That is,

$$Q_{t_n} = \sum_{m=1}^{n-1} \mathbf{q} \text{ for } n \geq 2,$$

and $Q_{t_n} = 0$ for $n = 0, 1$. We also call Q_{t_n} the period to date at time t_n. In addition, if we let Q_T be the total gas taken during the contract, then

$$\begin{aligned} Q_T &= \sum_{n=1}^{N-1} \mathbf{q} + q_{t_N} \\ &= Q_{t_N} + q_{t_N}. \end{aligned}$$

Note that Q_{t_N} is the period to date at time $t_N = T$, it does not include the volume of gas taken at time t_N. However, Q_T contains the volume of gas taken at time t_N.

- Then, upon taking the volume q_{t_n}, the payoff from the buyer's point of view at time t_n is equal to

$$q_{t_n}(S_n - I_n).$$

We call this payoff the instant payoff.

- Besides the daily exercise constraints (2.1), we also have global constraints on the total gas taken Q_T:

$$MB \leq Q_T \leq ACQ, \tag{2.2}$$

where MB and ACQ are also constants written in the contract. MB is called the minimum bill and ACQ is called the annual contract quantity.

Usually,

$$q_{\min} \cdot N < MB < ACQ \leq q_{\max} \cdot N. \tag{2.3}$$

Remark 2.1. In Edoli et al. (2013), the inequality (2.3) is called the non-trivial condition. They also suggest that this condition should be

$$q_{\min} \cdot N < MB < ACQ < q_{\max} \cdot N.$$

That is, if the buyer takes the daily maximum $q_{\max}$ on all t_n, $n = 1, \ldots, N$, then the resulting total consumption $q_{\max} \cdot N$ should be strictly larger than the annual contract quantity ACQ. There are contracts where $ACQ = q_{\max} \cdot N$ (see Breslin et al. (2008)), however, since it is natural that the gas supplier encourages the buyer to purchase more gas.

According to whether the global constraints (2.2) can be violated or not, GSAs can be divided into two types: GSAs with penalties and GSAs with firm constraints.

2.1.1 Gas sales agreements with firm constraints

In GSAs with firm constraints, the imposed constraints (2.2) cannot be violated. That is, any policy $\mathbf{q} = \{q_{t_n}\}_{n=1,2,\ldots,N}$ has to satisfy the following condition:

$$MB \leq \sum \mathbf{q} \leq ACQ.$$

We call this kind of policy the admissible policy and denote the collection of such policies by $\mathbf{q}_{\text{admis}}$. It means that the gas buyers may not have the access to the policy $\mathbf{q} = \{q_{t_n} = q_{\max}\}_{n=1,2,\ldots,N}$ or $\mathbf{q} = \{q_{t_n} = q_{\min}\}_{n=1,2,\ldots,N}$.

Through some simple calculations, we can obtain the policy:

$$\mathbf{q}_{\text{low}} = \{q_{t_n}^{\text{low}}\}_{n=1,2,\ldots,N}$$

which gives the lower bound of the period to date Q_{t_n} at each time step t_n, $n = 1, 2, \ldots, N$. That is, following $\mathbf{q}_{\text{low}}$, the buyer has the lowest period to date Q_{t_n} at each time t_n such that the period to date Q_{t_N} at time t_N equals $MB - q_{\max}$. The policy $\mathbf{q}_{\text{low}}$ can be obtained as follows:

- Let m_1 be the largest integer such that

$$m_1 \leq \frac{N \cdot q_{\max} - MB}{q_{\max} - q_{\min}} + 1 \tag{2.4}$$

- If $m_1 = 1$,
  - for $n = m_1 = 1$, set $q_{t_n}^{\text{low}} = MB - (N - m_1)q_{\max}$.
  - for $n = 2, \ldots, N$, set $q_{t_n}^{\text{low}} = q_{\max}$.
- If $1 < m_1 < N$,
  - for $n = 1, \ldots, m_1 - 1$, set $q_{t_n}^{\text{low}} = q_{\min}$.
  - for $n = m_1$, set $q_{t_n}^{\text{low}} = MB - (m_1 - 1)q_{\min} - (N - m_1)q_{\max}$.
  - for $n = m_1 + 1, \ldots, N$, set $q_{t_n}^{\text{low}} = q_{\max}$
- If $m_1 = N$,
  - for $n = 1, \ldots, m_1 - 1$, set $q_{t_n}^{\text{low}} = q_{\min}$.
  - for $n = m_1 = N$, set $q_{t_n}^{\text{low}} = MB - (m_1 - 1)q_{\min}$.
- If $m_1 > N$,
  - for $n = 1, \ldots, N$, set $q_{t_n}^{\text{low}} = q_{\min}$.

Remark 2.2. Indeed, the integer m_1 is obtained by finding the intersection between two lines: the line passing the point $(t_1, 0)$ with the slope of $q_{\min}$ and the line passing the point $(t_N, MB - q_{\max})$ with the slope of $q_{\max}$. Due to the fact that $q_{\min} < q_{\max}$ and the non-trivial condition (2.3), we have $N \cdot q_{\max} - MB > 0$ and $q_{\max} - q_{\min} > 0$. It follows that m_1 is an integer such that $m_1 \geq 1$. From (2.4), we have

$$\frac{N \cdot q_{\max} - MB}{q_{\max} - q_{\min}} \leq m_1 \leq \frac{N \cdot q_{\max} - MB}{q_{\max} - q_{\min}} + 1. \tag{2.5}$$

In addition, we have $q_{t_{m_1}}^{\text{low}}$ at time t_{m_1} which is given by

$$\begin{aligned} q_{t_{m_1}}^{\text{low}} &= MB - (m_1 - 1)q_{\min} - (N - m_1)q_{\max} \\ &= MB - N \cdot q_{\max} + q_{\min} + m_1(q_{\max} - q_{\min}). \end{aligned}$$

Together with (2.5), we have $q_{\min} \leq q_{t_{m_1}}^{\text{low}} \leq q_{\max}$.

Similarly, we can obtain the following policy:

$$\mathbf{q}_{\text{up}} = \{q_{t_n}^{\text{up}}\}_{n=1,2,\ldots,N}$$

which gives the upper bound of the period to date Q_{t_n} at each time step t_n, $n = 1, 2, \ldots, N$. Following $\mathbf{q}_{\text{up}}$, the buyer has the largest period to date Q_{t_n} at each time t_n such that the period to date Q_{t_N} at time t_N equals $ACQ - q_{\min}$. The policy $\mathbf{q}_{\text{up}}$ can be obtained as follows:

- Let m_2 be the largest integer such that

$$m_2 \leq \frac{ACQ - N \cdot q_{\min}}{q_{\max} - q_{\min}} + 1 \tag{2.6}$$

- If $m_2 = 1$,
  - For $n = m_2 = 1$, set $q_{t_n}^{\text{up}} = ACQ - (N - m_2)q_{\min}$.
  - For $n = m_2 + 1, \ldots, N$, set $q_{t_n}^{\text{up}} = q_{\min}$
- If $1 < m_2 < N$,
  - For $n = 1, \ldots, m_2 - 1$, set $q_{t_n}^{\text{up}} = q_{\max}$.
  - For $n = m_2$, set $q_{t_n}^{\text{up}} = ACQ - (m_2 - 1)q_{\max} - (N - m_2)q_{\min}$.
  - For $n = m_2 + 1, \ldots, N$, set $q_{t_n}^{\text{up}} = q_{\min}$
- If $m_2 = N$,
  - For $n = 1, \ldots, m_2 - 1$, set $q_{t_n}^{\text{up}} = q_{\max}$.
  - For $n = m_2 = N$, set $q_{t_n}^{\text{up}} = ACQ - (m_2 - 1)q_{\max}$.
- If $m_2 > N$,
  - For $n = 1, \ldots, N$, set $q_{t_n}^{\text{up}} = q_{\max}$.

Remark 2.3. The integer m_2 is obtained by finding the intersection between two lines: the line passing the point $(t_1, 0)$ with the slope of $q_{\max}$ and the line passing the point $(t_N, ACQ - q_{\min})$ with the slope of $q_{\min}$. Due to the non-trivial condition (2.3) and $q_{\min} < q_{\max}$, we have $ACQ - N \cdot q_{\min} > 0$ and $q_{\max} - q_{\min} > 0$. It follows that $m_2 \geq 1$. From (2.6), we have

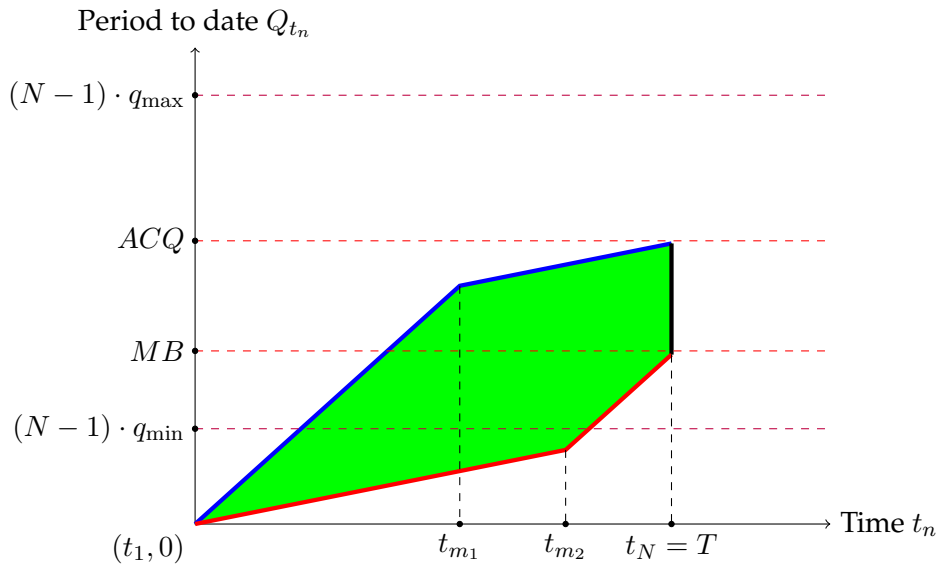
$$\frac{ACQ - N \cdot q_{\min}}{q_{\max} - q_{\min}} \leq m_2 \leq \frac{ACQ - N \cdot q_{\min}}{q_{\max} - q_{\min}} + 1. \tag{2.7}$$

In addition, we have $q^{\text{up}}_{t_{m_2}}$ at time t_{m_2} which is given by

$$\begin{aligned} q^{\text{up}}_{t_{m_2}} &= ACQ - (m_2 - 1)q_{\max} - (N - m_2)q_{\min} \\ &= ACQ - N \cdot q_{\min} + q_{\max} - m_2(q_{\max} - q_{\min}). \end{aligned}$$

Together with (2.7), we have $q_{\min} \leq q^{\text{up}}_{t_{m_2}} \leq q_{\max}$.

FIGURE 2.1: Admissible period to date areas the GSA with firm constraints



Given $\mathbf{q}_{\text{low}}$ and $\mathbf{q}_{\text{up}}$, the collection of the admissible period to date $\mathbf{Q}^{\text{admis}}_{t_n}$ at time t_n, $n = 2, \ldots, N$, is given by:

$$\mathbf{Q}^{\text{admis}}_{t_n} = \left\{ Q : \sum_{i=1}^{n-1} \mathbf{q}_{\text{low}} \leq Q \leq \sum_{i=1}^{n-1} \mathbf{q}_{\text{up}} \right\}.$$

In Figure 2.1, the green area shows the possible period to date at each time step. The lower bound policy, $\mathbf{q}_{\text{low}}$, and the upper bound policy, $\mathbf{q}_{\text{up}}$, can be obtained by following the red and blue lines, respectively.

We have the initial contract value at time 0, which is given by

$$V(0) = \sup_{\substack{q_{t_n} \in [q_{\min}, q_{\max}] \\ \sum_{n=1}^{N} q_{t_n} \in [MB, ACQ]}} \mathbb{E}\left[\sum_{n=1}^{N} e^{-t_n r} q_{t_n} \cdot (S_n - I_n) \right],$$

where $\mathbb{E}$ is the expectation with respect to risk-neutral probability measure $\mathbb{Q}$ which is the probability under which the gas price S has the dynamics. We call the decisions q_{t_n}, $n = 1, 2, \ldots, N$, which give the initial contract value $V(0)$, the optimal decisions.

It is suggested in Bardou, Bouthemy and Pagès (2009) (also see Bardou, Bouthemy and Pagès (2010)) that the GSA can be split into two parts: a swap part and a normalized GSA. This is called the decomposition of the GSA (also see Lari-Lavassani, Simchi and Ware (2001)). That is, the initial contract value $V(0)$ can be expressed in the following way:

$$V(0) = \underbrace{\mathbb{E}\left[\sum_{n=1}^{N} e^{-t_n r} q_{\min} \cdot (S_n - I_n)\right]}_{\text{The swap part}} + \underbrace{\sup_{\substack{q_{t_n} \in [0,\bar{q}] \\ \sum_{n=1}^{N} q_{t_n} \in [\bar{Q}_{\min}, \bar{Q}_{\max}]}} \mathbb{E}\left[\sum_{n=1}^{N} e^{-t_n r} q_{t_n} \cdot (S_n - I_n)\right]}_{\text{The normalized GSA}}. \tag{2.8}$$

The normalized GSA is a GSA contract with the following inputs: the daily minimum 0, the daily maximum $\bar{q} = q_{\max} - q_{\min}$, the minimum bill $\bar{Q}_{\min} = MB - N \cdot q_{\min}$ and the annual contract quantity $\bar{Q}_{\max} = ACQ - N \cdot q_{\min}$. The collection of the admissible policies of this normalized GSA $\mathbf{q}_{\text{admis}}^{[0,\bar{q}]} = \{q_{t_n}\}_{n=1,2,\ldots,N}$ is given by

$$\mathbf{q}_{\text{admis}}^{[0,\bar{q}]} = \left\{ q_{t_n}, n = 1, 2, \ldots, N \middle| \right. \\ \left. 0 \leq q_{t_n} \leq \bar{q}, n = 1, 2, \ldots, N \text{ and } \bar{Q}_{\min} \leq \sum_{n=1}^{N} q_{t_n} \leq \bar{Q}_{\max}. \right\}.$$

In Bardou, Bouthemy and Pagès (2010), the following theorem has been proved:

Theorem 2.1. *Consider the GSA formulated in Section 2.1.1, if the quantity*

$$\frac{ACQ - MB}{q_{max} - q_{min}}$$

is an integer, then the optimal exercise decision q_{t_n}, $n = 1, 2, \ldots, N$ *is either the daily maximum* $q_{\max}$ *or the daily minimum* $q_{\min}$.

This kind of consumption is called bang-bang consumption.

2.1.2 Gas sales agreements with penalties

In gas sales agreements with penalties, the constraints (2.2) can be violated. At the end of the contract, the total gas delivered under the GSA can be below the minimum bill or above the annual contract quantity. That is, the following scenarios are possible:

$$Q_T < MB \text{ or } Q_T > ACQ.$$

When (2.2) is violated, the contract holder has to pay penalties at the end of the contract, that is, at time $T = t_N$. The penalty functions are customized in different contracts, but they usually have the following form (see Barrera-Esteve et al. (2006)):

$$\mathcal{P}(x, Q_T) = -Ax \max\{MB - Q_T, 0\} - Bx \max\{Q_T - ACQ, 0\}, \tag{2.9}$$

where x can be the terminal gas price S_N, the terminal strike price I_N or a fixed constant. A and B are fixed constants written in the contract, and usually $A = B$. In this thesis, we let x be the terminal strike price I_N.

In addition, since constraints (2.2) can be violated, the gas buyer can purchase any quantity of gas q_{t_n} at each exercise date as long as it meets the daily exercise constraints $q_{\min} \leq q_{t_n} \leq q_{\max}$. The collection of the admissible policies $\mathbf{q}_{\text{admis}} = \{q_{t_n}\}_{n=1,2,\ldots,N}$ is given by:

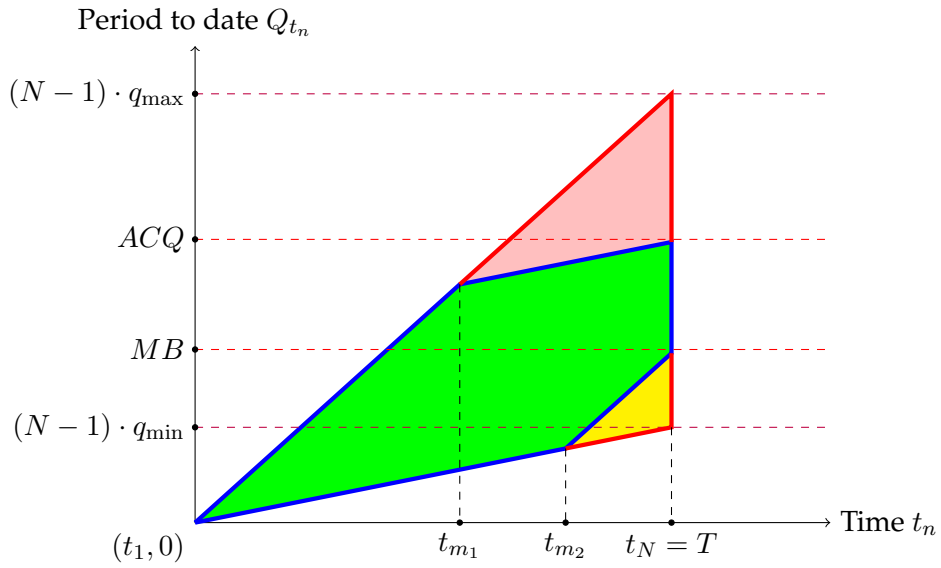
$$\mathbf{q}_{\text{admis}} = \{q_{t_n}, n = 1, 2, \ldots, N : q_{\min} \leq q_{t_n} \leq q_{\max}\}.$$

It follows that the collection of the admissible period to date $\mathbf{Q}_{t_n}^{\text{admis}}$ at time t_n, $n = 1, 2, \ldots, N$ is given by:

$$\mathbf{Q}_{t_n}^{\text{admis}} = \Big\{Q : (n-1)q_{\min} \leq Q \leq (n-1)q_{\max}\Big\}. \tag{2.10}$$

In Figure 2.2, the combination of the green, yellow and pink areas gives the possible period to date at each time step. If the period to date stays in the

FIGURE 2.2: Admissible period to date areas of the GSA with penalties



green area, then the buyer is free of penalties at the end of the contract. If the period to date lies in the pink area, then the buyer has to pay penalties because of the violation of the annual contract quantity. If the period to date is in the yellow area, then the buyer is obligated to pay penalties because of the violation of the minimum bill.

For such a GSA contract, we have the initial contract value:

$$V(0) = \sup_{q_{t_n} \in [q_{\min}, q_{\max}]} \mathbb{E}\left[\sum_{n=1}^{N} e^{-t_n r} q_{t_n} \cdot (S_n - I_n) + e^{-Tr}\mathcal{P}(I_N, Q_T)\right].$$

Similar to the GSA with firm constraints, we can also have the decomposition of the GSA with penalties. The initial value of the GSA with penalties can be

formulated in the following way:

$$V(0,S,I) = \underbrace{\mathbb{E}\left[\sum_{n=1}^{N} e^{-t_n r} q_{\min}\cdot(S_n - I_n)\right]}_{\text{The swap part}} + \underbrace{\sup_{\bar{q}_{t_n}\in[0,\bar{q}]} \mathbb{E}\left[\sum_{n=1}^{N} e^{-t_n r}\bar{q}_{t_n}\cdot(S_n - I_n) + e^{-Tr}\bar{\mathcal{P}}(I_N,\bar{Q}_T)\right]}_{\text{The normalized GSA}}. \quad (2.11)$$

This normalized GSA has the following inputs: the daily minimum 0, the daily maximum $\bar{q} = q_{\max} - q_{\min}$. The penalty function of this GSA is given by:

$$\bar{\mathcal{P}}(I_N,\bar{Q}_T) = \mathcal{P}(I_N,\bar{Q}_T + N\cdot q_{\min}),$$

where $\bar{Q}_T = \sum_{n=1}^{N}\bar{q}_{t_n}$.

Now, consider a GSA which has the penalty function $\mathcal{B}(I_N,Q_T) = I_N\cdot\mathfrak{P}(Q_T)$, where $\mathfrak{P}(Q)$ is a continuously differentiable function with respect to Q. In Barrera-Esteve et al. (2006), the following theorem has been proved:

Theorem 2.2. *Consider the GSA with the penalty function $\mathcal{B}(I_N,Q_T) = I_N\cdot\mathfrak{P}(Q_T)$. Suppose $\mathfrak{P}(Q)$ is continuously differentiable with respect to Q, and the following condition holds*

$$\mathbb{P}\left[e^{-rt_n}\Big(S_n - I_n\Big) + \mathbb{E}\Big(e^{-rT}\cdot I_N\mathfrak{P}'(Q_T^*)\Big|I_n,Q_{t_n}^*\Big) = 0\right] = 0,$$

where Q^* is the period to date obtained through optimal decisions, then the optimal decision at time t_n, $n = 1,2,\ldots,N-1$, is necessarily given by

$$q^*(t_n,S_n,I_n,Q_{t_n}^*) = q_{\max}\mathbf{1}_{\left\{e^{-rt_n}\left(S_n-I_n\right)+\mathbb{E}\left(e^{-rT}\cdot I_N\mathfrak{P}'(Q_T^*)|I_n,Q_{t_n}^*\right)>0\right\}} + q_{\min}\mathbf{1}_{\left\{e^{-rt_n}\left(S_n-I_n\right)+\mathbb{E}\left(e^{-rT}\cdot I_N\mathfrak{P}'(Q_T^*)|I(t_n),Q_{t_n}^*\right)<0\right\}}.$$

Remark 2.4. The above condition is hard to check since it involves the period to date Q^* which is obtained through optimal decisions. In addition, in our case, the penalty function is given by

$$\mathcal{P}(I_N,Q) = -I_N\cdot\max\big\{MB - Q, 0\big\} - I_N\cdot\max\big\{Q - ACQ, 0\big\},$$

which is clearly not differentiable at $Q = MB$ or $Q = ACQ$. That is, compared with Theorem 2.1, Theorem 2.2 is less useful since the penalty function is always not differentiable at some points in real contracts. For similar results which apply to the real contracts, we refer to Theorem 2.1(3) in Edoli et al. (2013) and Theorem 2 in Edoli and Vargiolu (2013). However, these theorems are derived for GSAs with firm constraints. Nonetheless, the bang-bang consumption is observed in our numerical examples.

Remark 2.5. As we can see in (2.8) and (2.11), the GSAs with firm constraints and the GSAs with penalties can both be decomposed to a swap part and a normalized GSA contract. In this thesis, we mainly focus on the part of the normalized GSA contract. Without loss of generality, we usually let $q_{\min} = 0$ and $q_{\max} = \bar{q}$, where $\bar{q}$ is a constant.

Remark 2.6. Based on Theorem 2.2, if we assume bang-bang consumption for the GSA with penalties and let $q_{\min} = 0$ and $q_{\max} = \bar{q}$, then the possible periods to date (2.10) at time t_n, $n = 1, \ldots, N$, are $0, \bar{q}, 2\bar{q}, \ldots, (n-1)\bar{q}$.

2.2 Construction of the two-dimensional trinomial tree

In this thesis, the pricing frameworks we build in Chapter 3 and Chapter 4 mainly evaluate the price of the gas sales agreement on a two-dimensional trinomial tree. In this section, we give a description on how we can construct a two-dimensional trinomial tree[1].

Suppose we have two energy products, E1 and E2. The log prices of E1 and E2 at time $t \in [0, T]$ are denoted by $X(t)$ and $Y(t)$, respectively. We assume that these log prices follow the mean-reverting processes

$$dX(t) = \Big[\theta^1(t) - \alpha_1 X(t)\Big]dt + \sigma_1 dB^1(t), \tag{2.12}$$

$$dY(t) = \Big[\theta^2(t) - \alpha_2 Y(t)\Big]dt + \sigma_2 dB^2(t), \tag{2.13}$$

where $B^1(t)$ and $B^2(t)$ are standard Brownian motions with correlation ρ, and α_1, σ_1, α_2 and σ_2 are constants. Let $F^1(t, T)$ and $F^2(t, T)$ be the forward prices of E1 and E2 at time $t \in [0, T]$ with the maturity T, respectively. $\theta^1(t)$

[1] This section contains essentially the content within Section 3 in the published paper Dong and Kang (2018).

and $\theta^2(t)$ are functions of time chosen to provide an exact fit to the observed forward curves $F^1(0,t)$ and $F^2(0,t)$. These processes are standard models when it comes to the commodity price in the energy market (see Hull (2011), Section 33.4).

In order to build a two-dimensional tree for (X,Y), we first build a simplified two-dimensional fundamental tree for a two-dimensional Markov process[2] (x,y) below, which is obtained by assuming $\theta^1(t) = \theta^2(t) = 0$ in (2.12) and (2.13):

$$dx(t) = -\alpha_1 x(t)dt + \sigma_1 dB^1(t) \tag{2.14}$$

$$dy(t) = -\alpha_2 y(t)dt + \sigma_2 dB^2(t) \tag{2.15}$$

Then we shift the nodes on the simplified tree by adding the correct drift in order to be consistent with the observed forward curve. In the rest of this section, we summarize the tree building procedures in Clewlow and Strickland (1999), Hull and White (1994a) and Hull and White (1994b).

Step one We build two separate fundamental one-dimensional trees for both x and y. Taking the lognormal price x as an example, the trinomial tree for y will be built in a very similar manner.

We discretize the time interval $[0,T]$ into N equally spaced time steps. That is, the time values are denoted by $t_n = n\triangle t$, where $n = 0,1,\ldots,N$ and $\triangle t = \frac{T}{N}$ is the time step. Similarly, the values of x at time t_n are referenced by $x_{n,s} = s\triangle X$, where s is an integer representing the level index and $\triangle X$ is the space step. This means that any node on the tree can be referenced by a pair of integers (n,s). In addition, due to convergence and stability considerations, it is suggested in Hull and White (1994a) that

$$\triangle X = \sigma_1\sqrt{3\triangle t}. \tag{2.16}$$

Because of the mean-reverting nature of this model, the trinomial tree will reach its maximum level, $s_{\max}$, at some point (the minimal level, $s_{\min}$, is also reached at the same time, $s_{\min} = -s_{\max}$). In the trinomial tree, the nodes emanating from node (n,s) are $(n+1,k+1)$, $(n+1,k)$ and $(n+1,k-1)$,

[2]The Markov property of this two-dimensional process can be checked by applying Theorem 7.1.2 in Øksendal (2003)

where k is chosen so that the value of x reached by the middle branch is as close as possible to the expected value of x at time t_{n+1}. Indeed,

$$k = \begin{cases} s-1 & s = s_{\max} \\ s & s_{\min} < s < s_{\max} \\ s+1 & s = s_{\min} \end{cases}. \tag{2.17}$$

We approximate (2.14) in discrete time as follows:

$$x(t_{n+1}) = x(t_n) - \alpha_1 x(t_n)\triangle t + \sigma_1\big(B^1(t_{n+1}) - B^1(t_n)\big). \tag{2.18}$$

Since $B^1(t_{n+1}) - B^1(t_n) \sim \mathcal{N}(0, \triangle t)$, together with (2.18), we would like the following properties to hold:

$$\mathbb{E}\Big[x(t_{n+1}) - x(t_n)\big|x(t_n) = x_{n,s}\Big] = -\alpha_1 x_{n,s}\triangle t, \tag{2.19}$$

$$\text{Var}\Big[x(t_{n+1}) - x(t_n)\big|x(t_n) = x_{n,s}\Big] = \sigma_1^2\triangle t.$$

It follows that

$$\mathbb{E}\Big[\big(x(t_{n+1}) - x(t_n)\big)^2\big|x(t_n) = x_{n,s}\Big] = \sigma_1^2\triangle t + \big(-\alpha_1 x_{n,s}\triangle t\big)^2. \tag{2.20}$$

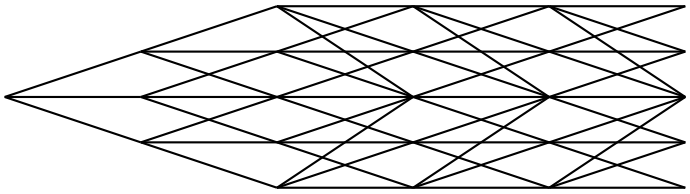
Let $p_{u,n,s}$, $p_{m,n,s}$ and $p_{d,n,s}$ be the probabilities associated with the upper, middle and lower branches emanating from node (n, s) respectively. Now, we find these probabilities in such a way as to satisfy (2.19) and (2.20):

$$\begin{aligned} -\alpha_1 x_{n,s}\triangle t =& \mathbb{E}\Big[x(t_{n+1}) - x(t_n)\big|x(t_n) = x_{n,s}\Big] \\ =& p_{u,n,s}((k+1) - s)\triangle X + p_{m,n,s}(k-s)\triangle X \\ &+ p_{d,n,s}((k-1) - s)\triangle X, \end{aligned} \tag{2.21}$$

and

$$\begin{aligned} \sigma_1^2\triangle t + (\alpha_1 x_{n,s}\triangle t)^2 =& \mathbb{E}\Big[\big(x(t_{n+1}) - x(t_n)\big)^2\big|x(t_n) = x_{n,s}\Big] \\ =& p_{u,n,s}((k+1) - s)^2\triangle X^2 + p_{m,n,s}(k-s)^2\triangle X^2 \\ &+ p_{d,n,s}((k-1) - s)^2\triangle X^2. \end{aligned} \tag{2.22}$$

FIGURE 2.3: A one-dimensional trinomial tree



Together with $p_{u,n,s} + p_{m,n,s} + p_{d,n,s} = 1$, we obtain

$$p_{u,n,s} = \frac{1}{2}\left[\frac{\sigma_1^2 \triangle t + (\alpha_1 x_{n,s} \triangle t)^2}{\triangle X^2} + (k-s)^2 - \frac{\alpha_1 x_{n,s} \triangle t}{\triangle X}(1 - 2(k-s)) - (k-s)\right], \tag{2.23}$$

$$p_{d,n,s} = \frac{1}{2}\left[\frac{\sigma_1^2 \triangle t + (\alpha_1 x_{n,s} \triangle t)^2}{\triangle X^2} + (k-s)^2 + \frac{\alpha_1 x_{n,s} \triangle t}{\triangle X}(1 + 2(k-s)) + (k-s)\right], \tag{2.24}$$

$$p_{m,n,s} = 1 - p_{u,n,s} - p_{d,n,s}. \tag{2.25}$$

It is shown in Hull and White (1994a) that, in order to ensure that $p_{u,n,s}$, $p_{m,n,s}$ and $p_{d,n,s}$ are all in the interval $[0, 1]$, $s_{\max}$ should be an integer between $\frac{0.184}{\alpha_1 \triangle t}$ and $\frac{0.816}{\alpha_1 \triangle t}$. To achieve the best efficiency, it is also suggested to set $s_{\max}$ at the smallest integer greater than $\frac{0.184}{\alpha_1 \triangle t}$. In the rest of this thesis, we will call this tree the fundamental tree. Figure 2.3 shows the structure of a fundamental tree with $s_{\max} = 2$.

Remark 2.7. Based on the above discussion, we can compute the probabilities by using (2.23), (2.24) and (2.25).

- If $s_{\min} < s < s_{\max}$, we have

$$\begin{cases} p_{u,n,s} = \frac{1}{6} + \frac{\alpha_1^2 s^2 \Delta t^2 - \alpha_1 s \Delta t}{2} \\ p_{m,n,s} = \frac{2}{3} - \alpha_1^2 s^2 \Delta t^2 \\ p_{d,n,s} = \frac{1}{6} + \frac{\alpha_1^2 s^2 \Delta t^2 + \alpha_1 s \Delta t}{2} \end{cases}. \tag{2.26}$$

- If $s = s_{\max}$, we have

$$\begin{cases} p_{u,n,s} = \frac{7}{6} + \frac{\alpha_1^2 s^2 \Delta t^2 - 3\alpha_1 s \Delta t}{2} \\ p_{m,n,s} = -\frac{1}{3} - \alpha_1^2 s^2 \Delta t^2 + 2\alpha_1 s \Delta t \\ p_{d,n,s} = \frac{1}{6} + \frac{\alpha_1^2 s^2 \Delta t^2 - \alpha_1 s \Delta t}{2} \end{cases} .$$

- If $s = s_{\min}$, we have

$$\begin{cases} p_{u,n,s} = \frac{1}{6} + \frac{\alpha_1^2 s^2 \Delta t^2 + \alpha_1 s \Delta t}{2} \\ p_{m,n,s} = -\frac{1}{3} - \alpha_1^2 s^2 \Delta t^2 - 2\alpha_1 s \Delta t \\ p_{d,n,s} = \frac{7}{6} + \frac{\alpha_1^2 s^2 \Delta t^2 + 3\alpha_1 s \Delta t}{2} \end{cases} .$$

Step two Once we have built two fundamental trinomial trees for both x and y, we combine these two trees together. Let node (n, z) be the node on the E2 tree at time t_n where the E2 price level is z. Any node on our new two-dimensional tree can be referenced by a triplet of integers (n, s, z), which indicates that the time is t_n, the level of the E1 price on the E1 tree is s, and the level of the E2 price on the E2 tree is z. There are three possible movements (an up movement, a middle movement and a down movement) at each node on both the E1 tree and the E2 tree, which gives a total of nine possible movements at each node on the two-dimensional tree. We let $\{\mathbf{m_1}, \mathbf{m_2}\}$ represent the movement on the two-dimensional tree, which is the combination of the $\mathbf{m_1}$ movement on the E1 tree and the $\mathbf{m_2}$ movement on the E2 tree. This means that we have the following nine movements with their corresponding probabilities:

$\{up, up\}$ with P_{uu}	$\{up, middle\}$ with P_{um}	$\{up, down\}$ with P_{ud}
$\{middle, up\}$ with P_{mu}	$\{middle, middle\}$ with P_{mm}	$\{middle, down\}$ with P_{md}
$\{down, up\}$ with P_{du}	$\{down, middle\}$ with P_{dm}	$\{down, down\}$ with P_{dd}

Denote the probabilities associated with the upper, middle and lower branches on the E1 tree by p_u, p_m and p_d, respectively. Similarly, denote the probabilities associated with the upper, middle and lower branches on the E2 tree by q_u, q_m and q_d, respectively. By assuming the correlation $\rho = 0$, the probabilities on the two-dimensional tree are the product of the corresponding probabilities associated with the branches on the E1 tree and the E2 tree. The

probability matrix is given by:

$$\Pi_{\rho=0} = \begin{pmatrix} P_{uu} & P_{um} & P_{ud} \\ P_{mu} & P_{mm} & P_{md} \\ P_{du} & P_{dm} & P_{dd} \end{pmatrix} = \begin{pmatrix} p_u q_u & p_u q_m & p_u q_d \\ p_m q_u & p_m q_m & p_m q_d \\ p_d q_u & p_d q_m & p_d q_d \end{pmatrix}. \tag{2.27}$$

When the correlation is non-zero, according to Appendix F in Brigo and Mercurio (2006), each probability is shifted to maintain the marginal distribution and the covariance structure of x and y. The probabilities are given by:

$$\Pi_{\rho>0} = \begin{pmatrix} P_{uu} & P_{um} & P_{ud} \\ P_{mu} & P_{mm} & P_{md} \\ P_{du} & P_{dm} & P_{dd} \end{pmatrix} = \begin{pmatrix} p_u q_u + 5\varepsilon & p_u q_m - 4\varepsilon & p_u q_d - \varepsilon \\ p_m q_u - 4\varepsilon & p_m q_m + 8\varepsilon & p_m q_d - 4\varepsilon \\ p_d q_u - \varepsilon & p_d q_m - 4\varepsilon & p_d q_d + 5\varepsilon \end{pmatrix} \tag{2.28}$$

and

$$\Pi_{\rho<0} = \begin{pmatrix} P_{uu} & P_{um} & P_{ud} \\ P_{mu} & P_{mm} & P_{md} \\ P_{du} & P_{dm} & P_{dd} \end{pmatrix} = \begin{pmatrix} p_u q_u - 1\varepsilon & p_u q_m - 4\varepsilon & p_u q_d + 5\varepsilon \\ p_m q_u - 4\varepsilon & p_m q_m + 8\varepsilon & p_m q_d - 4\varepsilon \\ p_d q_u + 5\varepsilon & p_d q_m - 4\varepsilon & p_d q_d - \varepsilon \end{pmatrix}, \tag{2.29}$$

where $\varepsilon = \frac{\rho}{36}$ if $\rho \geq 0$ and $\varepsilon = -\frac{\rho}{36}$ if $\rho < 0$.

Remark 2.8. Given any node (n, s, z) on the two-dimensional trinomial tree, (2.27), (2.28) and (2.29) can be calculated by using (2.23), (2.24) and (2.25). That is, given any node (n, s, z), we have $p_u = p_{u,n,s}$, $p_m = p_{m,n,s}$ and $p_d = p_{d,n,s}$. Same rule applies to q_u, q_m and q_d. It may happen that one or more entries in $\Pi_{\rho>0}$ or $\Pi_{\rho<0}$ are negative. To overcome this drawback, Brigo and Mercurio (2006) (Appendix F) suggest that one can substitute $\Pi_{\rho>0}$ or $\Pi_{\rho<0}$ with $\Pi_{\rho=0}$ on a node when negative probabilities happen on this node. That is, we assume different correlations on the nodes where some probabilities are negative. As reported in Brigo and Mercurio (2006), this action has a negligible impact on a derivative price as long as we have a sufficiently small Δt.

Step three Returning to the notations in Step one, we now make both the E1 tree and the E2 tree to be consistent with the observed forward curve by adding an amount, a_n, at each node of each time step t_n. Again, we use the E1 tree as an example, but the E2 tree can be adjusted in a very similar manner.

We define the state price, $G_{n,s}$, as the value of a security that pays 1 if the node (n, s) is reached, and 0 otherwise, at time 0. Then the state price at each

node can be obtained by forward induction:

$$G_{n+1,s} = \sum_{s'} G_{n,s'} p_{s',s} e^{-r\triangle t},$$

where $(n, s^{'})$ represents the node at time $n\triangle t$ that has a branch leading to node $(n+1, s)$ and $p_{s',s}$ is the probability of moving from node $(n, s^{'})$ to node $(n+1, s)$, r is the risk-free rate. Denote the E1 price by S, that is $S(t) = e^{X(t)}$. In addition, let $S_{n,s}$ be the E1 price where the time equals t_n and the level of the E1 price on the adjusted E1 tree is s. Thus, the price at time 0 of any European claim with payoff function $C(S_{n,s})$ at time step n on the tree is given by

$$C_0 = \sum_{s} G_{n,s} C(S_{n,s}), \tag{2.30}$$

where the summation takes place through all nodes (n, s) at time t_n. Now consider a case where $C(S_{n,s}) = S_{n,s}$, then according to (2.30), we have

$$e^{-rt_n} F^1(0, t_n) = \sum_{s} G_{n,s} S_{n,s}. \tag{2.31}$$

Let $X_{n,s} = x_{n,s} + a_n$. By letting $S_{n,s} = e^{x_{n,s}+a_n}$ and through equation (2.31), we obtain

$$e^{-rt_n} F^1(0, t_n) = \sum_{s} G_{n,s} e^{x_{n,s}+a_n}.$$

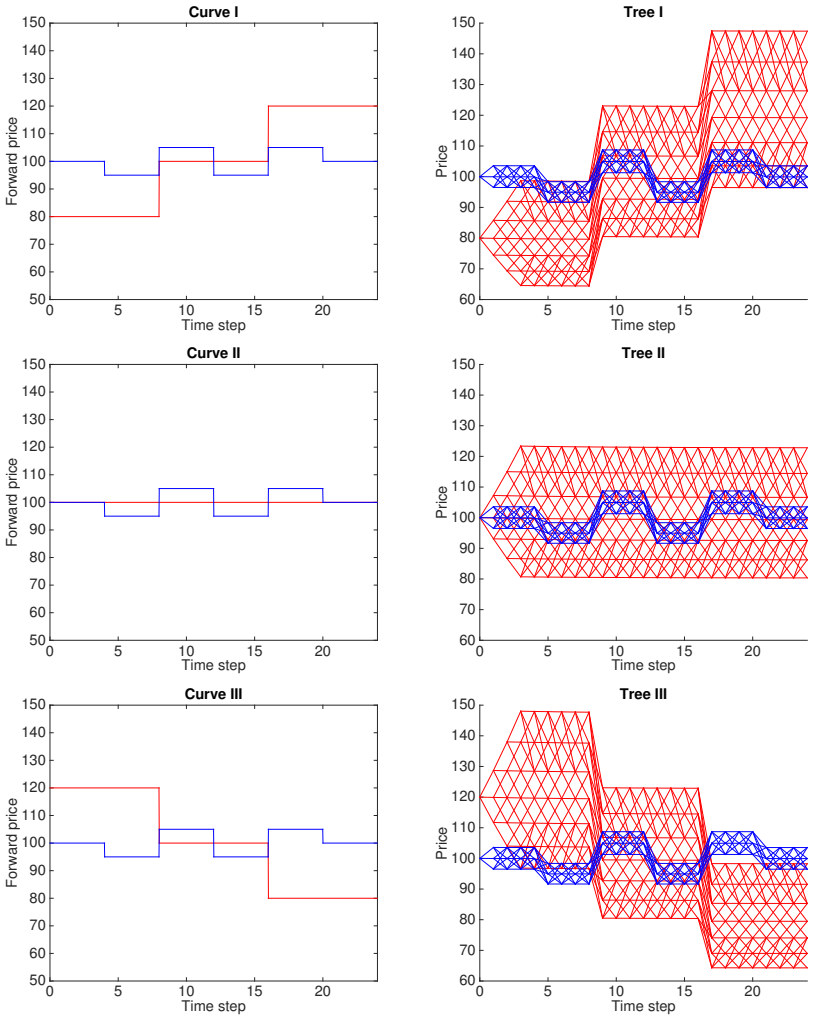
It follows that

$$a_n = \ln\left(\frac{e^{-rt_n} F^1(0, t_n)}{\sum_s G_{n,s} e^{x_{n,s}}}\right). \tag{2.32}$$

Therefore, by adding a_n at each node at each time t_n, we obtain the trinomial tree we need.

The right panel of Figure 2.4 demonstrates some examples of trinomial trees that have been constructed to be consistent with the forward curves shown in the left panel of Figure 2.4.

FIGURE 2.4: Trinomial tree examples



Note: The red plots refer to the E1 price and the blue plots refer to the E2 price. The parameters used are as follows: $T = 1$, $N = 24$, $\alpha_1 = 1.5$, $\alpha_2 = 5$, $\sigma_1 = 0.2$, $\sigma_2 = 0.1$ and $r = 0.05$.

Chapter 3

Analysis of a multiple year gas sales agreement

The main contents of this chapter have been published in Dong and Kang (2018).

3.1 Introduction

A gas sales agreement (GSA) is an American-style option with daily exercise opportunities. The delivery of daily quantities is constrained between pre-determined minimum and maximum daily limits. Under a GSA, the gas supplier encourages the buyer to purchase more gas when the market gas price is low, however, the buyer prefers to postpone the purchase until the market gas price is high. To protect the interests of gas suppliers, a minimum volume of gas, which is called the minimum bill, appears in the GSA. If the actual quantity of gas taken at each year end does not meet the minimum bill, the buyer will face a penalty. Typically, there is also a maximum annual quantity that can be taken. There are also two common and important features in multi-year GSA contracts: the make-up bank and the carry-forward bank. In the years where the gas taken is less than some reference level, the shortfall is added to a make-up bank. In later years, where the gas taken is greater than some reference level, the buyer can get some refund based on the quantity of extra gas purchased. In years where the gas taken is greater than some reference level, the excess gas is added to the carry-forward bank. In later years, the gas in the carry-forward bank can be used to reduce the minimum bill for that year. In addition, the amount that can be used from

the make-up and carry-forward banks each year is usually constrained by a recovery limit.

A number of papers have discussed the valuation of general swing options (GSAs without the make-up and carry-forward banks). Lattice-based numerical methods can be found in Thompson (1995), Jaillet, Ronn and Tompaidis (2004), Barrera-Esteve et al. (2006) and Bardou, Bouthemy and Pagès (2009) while simulation-based methods can be found in Dörr (2003), Meinshausen and Hambly (2004), Thanawalla (2006) and Ibáñez (2004). When the strike price of a GSA is a constant, the valuation of the GSA is a classical dynamic programming problem. In real contracts, however, the strike price is set based on the indexation principle under which the strike price is called the index. In each month, the value of the index is determined by the weighted average price of some energy substitutions in the previous month (see Asche, Osmundsen and Tveterås (2002) for details). This feature links the valuation of the GSAs to the moving average problem. So far, however, no effective method has been derived in the literature to value GSAs embedded with moving average features. In order to perform a detailed analysis of GSAs in terms of values, decisions, make-up usage, carry-forward usage, etc., we follow the assumption in Edoli et al. (2013) that the evaluation of the index follows a mean-reverting Markov process.

In cases where there are make-up and carry-forward banks, the GSA becomes even more complicated to evaluate. The basis features of both the make-up and carry-forward banks are explained in Breslin et al. (2008). In Holden, Løland and Lindqvist (2011), the authors propose an algorithm which uses the least-squares Monte Carlo method to evaluate GSAs with a carry-forward bank, although the authors only price the contract values and do not extract optimal decisions. The first formally quantitative treatment of GSAs containing the make-up provision was done in Edoli et al. (2013). The authors modelled one type of make-up clause and performed a sensitivity analysis of the contract value with respect to various parameters. The GSAs in Edoli et al. (2013) do not have penalties and the volume of gas in the make-up bank must be zero at the end of the contract. In addition, their GSAs contain either the make-up clause or the carry-forward clause, but not both. The second treatment of make-up and carry-forward banks appears in Chiarella, Clewlow and Kang (2016), where the authors evaluate GSAs

with both make-up and carry-forward banks in a regime-switching forward price curve model. The authors, however, only evaluate GSAs with constant strike prices and do not take the indexation into account. Furthermore, in both Chiarella, Clewlow and Kang (2016) and Edoli et al. (2013), the effect of the make-up and carry-forward banks on the optimal decisions is not given. The main contribution of this chapter, therefore, is to evaluate multiple year GSAs with both make-up and carry-forward banks, as in Breslin et al. (2008), but with stochastic strike prices (indexation), while also providing a detailed analysis of how the make-up bank, the carry-forward bank, the indexation and the different parameter settings affect both values and decisions. With the help of our numerical examples in Section 3.4, traders can have a better understanding of the impact of various important contract features on both the contract value and the trading strategies, such as recovery limits, penalties, the minimal bill, etc..

This chapter is organized as follows: In Section 3.2, we propose a mean-reverting model for the gas price and the index. Section 3.3 formulates both the cases of GSAs that do not have make-up and carry-forward banks and those that do. Section 3.4 compares GSAs with indexation with those with constant strike prices, and also analyzes the effect of make-up and carry-forward clauses, and how various parameters affect the contract value. At the end of Section 3.4 the ways in which the indexation affects decisions about weekly takes, make-up takes and carry-forward takes are analyzed. We draw conclusions in Section 3.5.

3.2 Modelling the gas price and the index

This chapter considers the one factor forward curve model built in Clewlow and Strickland (1999).[1] We assume the forward prices of both gas and index follow the stochastic differential equations below:

$$\frac{dF^S(t,T)}{F^S(t,T)} = \sigma_S e^{-\alpha_S(T-t)} dB^S(t), \tag{3.1}$$

[1]To model the whole forward price curve, a multi-factor model is usually required, but since we focus on computing the price and optimal decisions of the GSA, the overall volatility level matters more than the individual factors.

$$\frac{dF^I(t,T)}{F^I(t,T)} = \sigma_I e^{-\alpha_I(T-t)} dB^I(t), \tag{3.2}$$

where $F^S(t,T)$ and $F^I(t,T)$ are the forward prices of gas and index at time $t \in [0,T]$ with the maturity T, respectively. $B^S(t)$ and $B^I(t)$ are standard Brownian motions with correlation ρ, and α_S, σ_S, α_I and σ_I are constants.

We now proceed to get the spot and log prices of the gas and index in analogy with Clewlow and Strickland (1999). Given (3.1), we have the following solution:

$$F^S(t,T) = F^S(0,T) \exp\left[-\frac{1}{2}\sigma_S^2 \int_0^t e^{-2\alpha_S(T-u)} du + \sigma_S \int_0^t e^{-\alpha_S(T-u)} dB^S(u)\right].$$

Then the spot price is obtained by setting $T = t$:

$$S(t) = F^S(0,t) \exp\left[-\frac{1}{2}\sigma_S^2 \int_0^t e^{-2\alpha_S(t-u)} du + \sigma_S \int_0^t e^{-\alpha_S(t-u)} dB^S(u)\right]. \tag{3.3}$$

According to Appendix A in Clewlow and Strickland (1999) (equations (A.3) and (A.4)), we have the following equation:

$$\begin{aligned}\frac{dS(t)}{S(t)} =& \left[\frac{\partial \ln F^S(0,t)}{\partial t} + \alpha_S \sigma_S^2 \int_0^t e^{-2\alpha_S(t-u)} du \right.\\ & \left. - \alpha_S \sigma_S \int_0^t e^{\alpha_S(t-u)} dB^S(u)\right] dt + \sigma_S dB^S(t).\end{aligned}$$

By (3.3), we have

$$\ln S(t) = \ln F^S(0,t) - \frac{1}{2}\sigma_S^2 \int_0^t e^{-2\alpha_S(t-u)} du + \sigma_S \int_0^t e^{-\alpha_S(t-u)} dB^S(u),$$

which gives

$$\alpha_S \sigma_S \int_0^t e^{\alpha_S(t-u)} dB^S(u) = \alpha_S \left[\ln S(t) - \ln F^S(0,t) + \frac{1}{2}\sigma_S^2 \int_0^t e^{-2\alpha_S(t-u)} du\right].$$

By the fact that

$$\int_0^t e^{-2\alpha_S(t-u)} du = \frac{1 - e^{-2\alpha_S t}}{2\alpha_S},$$

we have

$$\frac{dS(t)}{S(t)} = \left[\frac{\partial \ln F^S(0,t)}{\partial t} + \alpha_S(\ln F^S(0,t) - \ln S(t)) + \frac{\sigma_S^2(1-e^{-2\alpha_S t})}{4}\right] dt + \sigma_S dB^S(t).$$

After applying Itô's formula with $X(t) = \ln S(t)$, we get

$$dX(t) = \phi^S(t, X(t))dt + \sigma_S dB^S(t), \tag{3.4}$$

where

$$\phi^S(t,x) = \frac{\partial \ln F^S(0,t)}{\partial t} + \alpha_S(\ln F^S(0,t) - x) + \frac{\sigma_S^2}{4}(1-e^{-2\alpha_S t}) - \frac{1}{2}\sigma_S^2.$$

Through similar calculations, with $Y(t) = \ln I(t)$, for (3.2) we have

$$dY(t) = \phi^I(t, Y(t))dt + \sigma_I dB^I(t),$$

where

$$\phi^I(t,y) = \frac{\partial \ln F^I(0,t)}{\partial t} + \alpha_I(\ln F^I(0,t) - y) + \frac{\sigma_I^2}{4}(1-e^{-2\alpha_I t}) - \frac{1}{2}\sigma_I^2.$$

To check the Markov property of the two-dimensional process (X, Y), we rewrite X and Y. The Brownian motions B^S and B^I can be expressed in terms of two independent Brownian motions B_1 and B_2, where $B^S = B_1$ and $B^I = \rho B_1 + \sqrt{1-\rho^2}B_2$. Then (X, Y) can be written as follows:

$$d\begin{pmatrix} X(t) \\ Y(t) \end{pmatrix} = A_{2\times 3}\begin{pmatrix} dt \\ dB_1 \\ dB_2 \end{pmatrix},$$

where

$$A_{2\times 3} = \begin{pmatrix} \phi^S(t, X(t)) & \sigma_S & 0 \\ \phi^I(t, Y(t)) & \sigma_I\rho & \sigma_I\sqrt{1-\rho^2} \end{pmatrix}.$$

Being each entry of matrix $A_{2\times 3}$ a Lipschitz function and $(t, B_1(t), B_2(t))$ a set of independent Lévy processes, one can apply Theorem 32 of Chapter V in Protter (2005) to confirm the Markov property of (X, Y).

To use the tree building procedures in Section 2.2, we rewrite the log prices of the gas and index as follows:

$$dX(t) = \left[\theta^S(t) - \alpha_S X(t)\right] dt + \sigma_S dB^S(t), \tag{3.5}$$

$$dY(t) = \left[\theta^I(t) - \alpha_I Y(t)\right] dt + \sigma_I dB^I(t), \tag{3.6}$$

where

$$\begin{aligned}\theta^S(t) =& \phi^S(t, X(t)) + \alpha_S X(t) \\ =& \frac{\partial \ln F^S(0,t)}{\partial t} + \alpha_S \ln F^S(0,t) + \frac{\sigma_S^2}{4}(1 - e^{-2\alpha_S t}) - \frac{1}{2}\sigma_S^2,\end{aligned}$$

and

$$\begin{aligned}\theta^I(t) =& \phi^I(t, Y(t)) + \alpha_I Y(t) \\ =& \frac{\partial \ln F^I(0,t)}{\partial t} + \alpha_I \ln F^I(0,t) + \frac{\sigma_I^2}{4}(1 - e^{-2\alpha_I t}) - \frac{1}{2}\sigma_I^2.\end{aligned}$$

3.3 Multiple year gas sales agreements

As introduced in Section 3.1, the GSA is an American-style option between a gas supplier and a gas buyer for the delivery of daily quantities of gas at variable strike prices over several years. In the absence of make-up and carry-forward banks, the main constraint in a GSA is the minimum bill. In each year, a penalty is applied if the actual gas taken is below the minimum bill. When it comes to the make-up and carry-forward banks, however, the contract becomes more complicated. In the years where the gas taken is less than the minimum bill, the shortfall is added to a make-up bank. In years where the gas taken is greater than some reference level (this level is called the carry-forward base), the excess gas is added to the carry-forward bank. In later years, under some pre-specified conditions, the gas in the make-up bank can be withdrawn to get a refund, while the gas in the carry-forward bank can be withdrawn to reduce the minimum bill. Since both the make-up and carry-forward banks offer the buyer opportunities to reduce possible losses, some pre-specified limits are imposed to protect the interests of the gas supplier. For example, in each year, the gas withdrawn from these banks cannot exceed some pre-specified limits: the carry-forward bank recovery limit and

the make-up bank recovery limit. These recovery limits will play large roles in the contract value. Also, since both the make-up and carry-forward banks are related to the minimum bill (and hence related to the penalty), how the penalty is calculated will also greatly affect the contract value. In addition, in each year, the evaluation of the make-up and carry-forward banks should be formulated carefully.

In this section, we provide an analytical representation of GSA contracts with the above features. GSA contracts without make-up and carry-forward banks are presented in Subsection 3.3.1. GSA contracts with make-up and carry-forward banks are formulated in Subsection 3.3.2.

3.3.1 Gas sales agreements in the absence of make-up and carry-forward banks

A gas sales agreement with indexation between the gas provider and the gas user could have many specific features to satisfy the needs of both parties to the agreement but these contracts usually have a number of common features as set out below:

- The contract lasts L years and there are J time periods (typically days) in each year, L and J are positive integers. Let T_i be the end of each year i, for $i = 1, \ldots, L$. Obviously, T_L is the maturity of this contract.

- Let the time interval be $\triangle t = 1/J$, then the contract duration is equally spaced into $J \cdot L$ periods. Denote the j-th period of the i-th year by $[t_{i,j-1}, t_{i,j}]$, $i = 1, \ldots, L$ and $j = 0, \ldots, J$, where $t_{i,j} = (i-1) + j \cdot \Delta t$. Hence we have

$$0 = t_{1,0} < t_{1,1} < \ldots < t_{1,J} = T_1 = t_{2,0} < t_{2,1} < \ldots$$
$$\ldots < t_{L-1,J} = T_{L-1} = t_{L,0} < t_{L,1} < \ldots < t_{L,J} = T_L.$$

  We assume that the holder of a GSA has exactly one exercise opportunity at each time $t_{i,j}$ in each gas year where $i = 1, \ldots, L$ and $j = 1, \ldots, J$, which amounts to J rights for a whole year. Typical GSAs can usually be exercised daily ($J = 365$) or weekly ($J = 52$).

- Let $q_{t_{i,j}}$ be the amount of gas taken (exercise decision) at time $t_{i,j}$, which is constrained by the minimum daily quantity $q_{\min}$ and the maximum

daily quantity $q_{\max}$, that is $q_{\min} \leq q_{t_{i,j}} \leq q_{\max}$. Let $\mathbf{q}_{\text{admis}}$ be the collection of admissible policies which is given by

$$\mathbf{q}_{\text{admis}} = \left\{q_{t_{i,j}}, i = 1, \ldots, L \text{ and } j = 1, \ldots, J : q_{\min} \leq q_{t_{i,j}} \leq q_{\max}\right\}.$$

Let $Q_{t_{i,j}}$ be the cumulative amount of gas taken before time $t_{i,j}$ in year i (also known as the period to date). It is straightforward to see that

$$Q_{t_{i,j}} = \sum_{k=1}^{j-1} q_{t_{i,k}}. \tag{3.7}$$

In addition, we let Q_{T_i} be the total amount of gas taken during the year i. That is,

$$Q_{T_i} = Q_{t_{i,J}} + q_{t_{i,J}}. \tag{3.8}$$

- Let $S_{i,j}$ and $I_{i,j}$ be the gas price and the index at time $t_{i,j}$, respectively. Then, upon taking the volume $q_{t_{i,j}}$, the payoff from the buyer's point of view at time $t_{i,j}$ is
$$q_{t_{i,j}} \left(S_{i,j} - I_{i,j}\right).$$

- In each year, there is a maximum quantity of gas the buyer can take, which is called the annual contract quantity. Denote the annual contract quantity in year i by ACQ_i. In each year, the amount of gas the buyer has taken cannot exceed the annual contract quantity. That is, $Q_{T_i} \leq ACQ_i$ for each year i. Similarly, there is a minimum quantity of gas the buyer has to take in each year, which is called the minimum bill. Denote the minimum bill in year i by MB_i. The total gas taken in each year can be below the minimum bill, however, in which case the buyer has to pay penalties at the end of that year. More precisely, if $Q_{T_i} < MB_i$ in year i, at the end of year i, there is an out cash flow generated by the penalty, in addition to the cash flow generated by the exercise decision.

- Usually, the penalty rate is a percentage of the index price. In this chapter, we let the penalty be given by
$$\eta \cdot I_{i,J} \cdot \min\left\{Q_{T_i} - MB_i, 0\right\}$$

for each year i, where $\eta \in [0,1]$ is a constant, which is called the penalty coefficient.

- The risk-free rate is denoted by r.

From the buyer's point of view, the goal is to maximize the total expected discounted payoff of the contract, including the penalty. That is, to find the value V_0 of this contract at time 0 which is given by

$$V_0 = \sup_{q_{t_{i,j}} \in [q_{\min}, q_{\max}]} \mathbb{E}\left[\sum_{i=1}^{L}\left(\sum_{j=1}^{J} e^{-rt_{i,j}} q_{t_{i,j}}(S_{i,j} - I_{i,j}) + e^{-rt_{i,J}} \cdot \eta \cdot I_{i,J} \cdot \min\{Q_{T_i} - MB_i, 0\}\right)\right]. \quad (3.9)$$

Remark 3.1. In Edoli et al. (2013), the author claims that the non-trivial constraints below are extremely important:

$$q_{\min} \cdot J < MB_i \leq Q_{T_i} \leq ACQ_i < q_{\max} \cdot J \text{ for each } i.$$

Since under normal circumstances, however, the gas supplier would encourage the buyer to purchase more gas, it is natural to have $ACQ_i = q_{\max} \cdot J$, namely

$$q_{\min} \cdot J < MB_i \leq Q_{T_i} \leq ACQ_i = q_{\max} \cdot J \text{ for each } i.$$

In a GSA contract, the user and the supplier can specify an existing volume of gas which has been taken at the beginning of each gas year. That is, if the gas supplier would like to offer the gas buyer a discount in a specific year, he/she can assume that there is an existing volume of gas has been taken at the beginning of this year. We denote this volume of gas by $\tilde{Q}_i$. In this thesis, we let $\tilde{Q}_i = 0, \forall i$, but it is straightforward to relax this assumption as long as we know the exact value specified in a particular contract. In addition, we further assume that, once the annual contract quantity is met, the daily minimum quantity can be violated, since the amount of gas the buyer has taken cannot exceed the annual contract quantity. That is, we let $q_{t_{i,j}} = 0$ as long as $Q_{t_{i,j}} = ACQ_i$, and we omit the discussion of this scenario in the rest of this chapter.

Dynamic Programming On the last day of the contract $t_{L,J}$, we have the following three scenarios: if $S_{L,J} > I_{L,J}$, that is the payoff is increasing in the volume purchased then we should take as much gas as we can; if $S_{L,J} \leq I_{L,J}$ and $I_{L,J} - S_{L,J} < \eta \cdot I_{L,J}$, we can reduce the total loss by purchasing a quantity up to that required to avoid the penalty, or the maximum possible, whichever is smaller; if $S_{L,J} \leq I_{L,J}$ and $I_{L,J} - S_{L,J} \geq \eta \cdot I_{L,J}$, that is the loss on the purchase of the gas is not compensated by the reduction in the penalty payment, we take as little gas as we can.

More precisely, let $q^* = q^*(S, I, Q, t)$ be the optimal take decision where the gas price equals S, the index equals I and the period to date equals Q at time t. In particular, on the last day $t_{L,J}$ of the contract, we have:

$$q^* = \begin{cases} q_{\max}, & S > I, \\ \min\Big\{ \max\big\{MB_L - Q, q_{\min}\big\}, q_{\max}\Big\}, & S \leq I \text{ and } (1-\eta)I < S, \\ q_{\min}, & S \leq I \text{ and } (1-\eta)I \geq S, \end{cases}$$

On the final day, the buyer still has to face the penalty if the minimum bill is not reached. After the last decision $q^*(S, I, Q, t_{L,J})$, the total amount of gas taken in the last year is known. Then the penalty becomes:

$$\eta \cdot I_{L,J} \cdot \min\Big\{Q_{t_{L,J}} + q^*(S, I, Q, t_{L,J}) - MB_L, 0\Big\}.$$

Denote the value of the contract by $V(S, I, Q, t)$ where the gas price equals S, the index equals I and the period to date equals Q at time t. The terminal payoff $V(S, I, Q, t_{L,J})$ is determined by the instant strike payoff and the penalty, which leads to the following expression:

$$\begin{aligned} &V(S, I, Q, t_{L,J}) \\ =&q^*(S, I, Q, t_{L,J}) \cdot (S - I) + \eta \cdot I \cdot \min\Big\{0, Q + q^*(S, I, Q, t_{L,J}) - MB_L\Big\}. \end{aligned}$$

Now that we have the terminal payoff, we can work backwards in time to find the optimal take and optimal value through the life of the contract. In fact, at each day $t_{i,j}$ within a year, we should choose the optimal exercise

decision according to

$$\begin{aligned}
&q^*(S,I,Q,t_{i,j})\\
=&\text{argmax}_{q_{t_{i,j}}\in[q_{\min},q_{\max}]}\Big\{q_{t_{i,j}}(S-I)+\\
&\mathbb{E}\left[e^{-r\triangle t}V(S_{i,j+1},I_{i,j+1},Q+q_{t_{i,j}},t_{i,j+1})\Big|S_{i,j}=S,I_{i,j}=I\right]\Big\},
\end{aligned}$$

and the contract value is

$$\begin{aligned}
&V(S,I,Q,t_{i,j})\\
=&q^*(S,I,Q,t_{i,j})\cdot(S-I)+\\
&\mathbb{E}\left[e^{-r\triangle t}V(S_{i,j+1},I_{i,j+1},Q_{t_{i,j}}+q^*(S,I,Q,t_{i,j}),t_{i,j+1})\mid S_{i,j}=S,I_{i,j}=I\right].
\end{aligned}$$

On the last day of each year (except the final year of the contract), we choose the optimal exercise decision according to

$$\begin{aligned}
&q^*(S,I,Q,t_{i,J})\\
=&\text{argmax}_{q_{t_{i,J}}\in[q_{\min},q_{\max}]}\Big\{q_{t_{i,J}}(S-I)+\eta\cdot I\cdot\min\Big\{0,Q+q_{t_{i,j}}-MB_i\Big\}+\\
&\mathbb{E}\left[e^{-r\triangle t}V(S_{i+1,1},I_{i+1,1},\tilde{Q}_i,t_{i+1,1})\mid S_{i,J}=S,I_{i,J}=I\right]\Big\},
\end{aligned}\tag{3.10}$$

where $\tilde{Q}_i$ is the existing volume of gas which has been taken at the beginning of year i (recall Remark 3.1). Then the contract value is obtained by

$$\begin{aligned}
&V(S,I,Q,t_{i,J})\\
=&q^*(S,I,Q,t_{i,J})\cdot(S-I)+\eta\cdot I\cdot\min\Big\{0,Q+q^*(S,I,Q,t_{i,J})-MB_i\Big\}+\\
&\mathbb{E}\left[e^{-r\triangle t}V(S_{i+1,1},I_{i+1,1},\tilde{Q}_i,t_{i+1,1})\mid S_{i,J}=S,I_{i,J}=I\right].
\end{aligned}\tag{3.11}$$

3.3.2 Multi-year gas sales agreements with both make-up and carry-forward banks

Turning now to the make-up and carry-forward features of multiple year GSA contracts, which were previewed in Section 3.1 of this chapter, we now formulate these two banks based on the notations in Subsection 3.3.1 above:

- Denote the amount of gas available in the carry-forward bank in year i by C_i, $i=1,\ldots,L$. Let CRL_i be the carry-forward bank recovery limit

and CB_i be the carry-forward base in year i. In addition, denote the usage of gas in the carry-forward bank by c_i in year i. Denote the amount of gas available in the make-up bank in year i by M_i, $i = 1, \ldots, L$. Let MRL_i be the make-up bank recovery limit and m_i be the usage of gas in the make-up bank in year i.

- In year i, when the total gas taken is $Q_{T_i} > \max\{MB_i + m_i, CB_i\}$, then the excess gas is added to the carry-forward bank, which gives $C_{i+1} = C_i + (Q_{T_i} - \max\{MB_i + m_i, CB_i\})$. In year i, if the total gas taken is $Q_{T_i} < MB_i$, the buyer can use the gas in the carry-forward bank to reduce the minimum bill, which adjusts the minimum bill in year i to $MB_i - c_i$. In addition, the possible year end penalty becomes $\eta \cdot I_{i,J} \cdot \min\{Q_{T_i} - (MB_i - c_i), 0\}$. The usage c_i, however, is constrained by the amount of gas in the carry-forward bank and the recovery limit, that is, $c_i \leq \min\{C_i, CRL_i\}$. Based on the above formulation, we obtain the evolution of the carry-forward bank, which is given by

$$C_{i+1} = (C_i - c_i) + \max\Big\{Q_{T_i} - \max\big\{MB_i + m_i, CB_i\big\}, 0\Big\}.$$

- In year i, when the gas taken is less than the adjusted minimum bill, that is $Q_{T_i} < MB_i - c_i$, the shortfall is added to the make-up bank and $M_{i+1} = M_i + (MB_i - c_i - Q_{T_i})$. In year i, when the gas taken is $Q_{T_i} > MB_i + C_i$, a refund will be given based on the excess gas. Then at the end of those years, in addition to the instant exercise payoff, an income cashflow is generated from the refund, which is given by $I_{i,J} \cdot m_i$; noting, of course, that $m_i \leq \min\{M_i, MRL_i, \max\{Q_{T_i} - (MB_i + C_i), 0\}\}$. We can also obtain the evolution of the make-up bank by

$$M_{i+1} = (M_i - m_i) + \max\Big\{MB_i - c_i - Q_{T_i}, 0\Big\}.$$

The objective function With both make-up and carry-forward banks, the payoff at the end of each year not only depends on the instant exercise and the possible penalty, but also the possible income from the refund. That is, at the end of each year i, we have the following cash flows:

- the instant exercise payoff $q_{t_{i,J}}\,(S_{i,J} - I_{i,J})$.

- the possible penalty if the total gas taken is less than the adjusted minimum bill $MB_i - c_i$:

$$\eta \cdot I_{i,J} \cdot \min\Big\{Q_{t_{i,J}} + q_{t_{i,J}} - (MB_i - c_i), 0\Big\}.$$

- the income generated by the possible refund when the total gas taken exceeds $MB_i + C_i$:

$$I_{i,J} \cdot \min\Big\{m_i, \max\big\{Q_{t_{i,J}} + q_{t_{i,J}} - (MB_i + C_i), 0\big\}\Big\}.$$

The payoff for each day within a year is the same as that with GSAs in the absence of make-up and carry-forward banks. Thus, from the perspective of the buyer, the goal is to find the initial contract value V_0 that is given by

$$\begin{aligned} V_0 = \sup_{q_{t_{i,j}}, m_i, c_i} \mathbb{E}\Bigg[\sum_{i=1}^{L}\Bigg(\sum_{j=1}^{J} e^{-rt_{i,j}} q_{t_{i,j}}(S_{i,j} - I_{i,j}) \\ + e^{-rt_{i,J}} \cdot \eta \cdot I_{i,J} \cdot \min\Big\{Q_{t_{i,J}} + q_{t_{i,J}} - (MB_i - c_i), 0\Big\} \\ + e^{-rt_{i,J}} \cdot I_{i,J} \cdot \min\Big\{m_i, \max\Big\{Q_{t_{i,J}} + q_{t_{i,J}} - (MB_i + C_i), 0\Big\}\Big\}\Bigg)\Bigg]. \end{aligned} \tag{3.12}$$

Remark 3.2. Recall (2.8) and (2.11) in Section 2.1, we have shown that a one-year GSA can be split into two parts: a swap part and a normalized GSA. The multiple years GSA can also be split into these two parts. Rewrite (3.12) as follows:

$$\begin{aligned} V_0 =& \mathbb{E}\Bigg[\sum_{i=1}^{L}\Bigg(\sum_{j=1}^{J} e^{-rt_{i,j}} q_{\min}(S_{i,j} - I_{i,j})\Bigg)\Bigg] \\ &+ \sup_{\bar{q}_{t_{i,j}}, m_i, c_i} \mathbb{E}\Bigg[\sum_{i=1}^{L}\Bigg(\sum_{j=1}^{J} e^{-rt_{i,j}} \bar{q}_{t_{i,j}}(S_{i,j} - I_{i,j}) \\ &+ e^{-rt_{i,J}} \cdot \eta \cdot I_{i,J} \cdot \min\Big\{Q_{t_{i,J}} + \bar{q}_{t_{i,J}} + q_{\min} - (MB_i - c_i), 0\Big\} \\ &+ e^{-rt_{i,J}} \cdot I_{i,J} \cdot \min\Big\{m_i, \max\Big\{Q_{t_{i,J}} + \bar{q}_{t_{i,J}} + q_{\min} - (MB_i + C_i), 0\Big\}\Big\}\Bigg)\Bigg], \end{aligned} \tag{3.13}$$

where $\bar{q}_{t_{i,j}} \in [0, q_{\max} - q_{\min}]$,

$$Q_{t_{i,j}} = \sum_{k=1}^{j-1} \left(\bar{q}_{t_{i,j}} + q_{\min} \right), \tag{3.14}$$

$i = 1, \ldots, L$, $j = 2, \ldots, J$ and $Q_{t_{i,1}} = 0$. The first expectation in (3.13) refers to the swap part, and the rest of (3.13) refers to the normalized GSA. This normalized GSA holds basically the same characteristics with the multi-year GSA introduced in this subsection except two differences. The first difference is that this normalized GSA uses (3.14) to calculate the period to date $Q_{t_{i,j}}$ instead of using (3.7). The second difference is that this normalized GSA calculate the total amount of gas taken during the year $i = 1, \ldots, L$ through

$$Q_{T_i} = Q_{t_{i,J}} + \bar{q}_{t_{i,J}} + q_{\min}$$

instead of using (3.8). Since the multi-year GSA also have the so-called decomposition of the GSA, without loss of generality, we can let $q_{\min} = 0$ and $q_{\max} = 1$.

The terminal condition With both make-up and carry-forward banks, the decision to be made on the last day is much more complicated. On the last day of the contract, we need to use as much of the balance in both the make-up and carry-forward banks as possible. Let $q^*(S, I, Q, C, M, t)$ and $V(S, I, Q, C, M, t)$ be the optimal exercise decision and the contract value respectively where the gas price equals S, the index equals I, the period to date equals Q, the amount of gas in the carry-forward bank equals C and the amount of gas in the make-up bank equals M, at time t. Next, let us find $q^*(S, I, Q, C, M, t_{L,J})$. On the last day, if $S > I$, that is the payoff is increasing in the volume purchased, then we should take as much gas as we can. If $S \leq I$, however, we have the following scenarios:

- If there is no gas available in the make-up bank, or the buyer has to take more gas to meet the minimum bill, that is $M = 0$ or $Q < MB_L$.
  - If $(1 - \eta) I < S$, that is the loss on the purchase of the gas is compensated by the reduction in the penalty payment, then the optimal choice is to purchase a quantity up to the amount required to

avoid the penalty, or the maximum possible, whichever is smaller. It follows that

$$q^*\left(S,I,Q,C,M,t_{L,J}\right)=\min\Big\{\max\left\{MB_L-C-Q,q_{\min}\right\},q_{\max}\Big\}.$$

– If $(1-\eta)\,I\geq S$, that is the loss on the purchase of the gas is not compensated by the reduction in the penalty payment, then

$$q^*\left(S,I,Q,C,M,t_{L,J}\right)=q_{\min}.$$

- If there is gas available in the make-up bank, that is $M>0$, and $Q\geq MB_L+C$, this means that the buyer can get a refund from the make-up bank.

  – If there is gas available in the make-up bank, that is $0\leq Q-(MB_L+C)<M$, the buyer should buy a quantity of gas up to the gas available in the make-up bank, or the maximum possible:

$$\begin{aligned}&q^*(S,I,Q,C,M,t_{L,J})\\&=\max\Big\{\min\left\{M-(Q-(MB_L+C)),q_{\max}\right\},q_{\min}\Big\}.\end{aligned}$$

  – If the buyer cannot take more refund because of a shortage of gas in the make-up bank, that is $Q-(MB_L+C)\geq M$, then they should take the daily minimum so as to avoid the loss:

$$q^*\left(S,I,Q,C,M,t_{L,J}\right)=q_{\min}.$$

- If $M>0$, but the period to date is not yet large enough to get a refund, that is $MB_L\leq Q<(MB_L+C)$.

  – If $(MB_L+C)-Q<q_{\max}$, the buyer can get some refund by taking $q_{\max}$. In this scenario, if the buyer decides to get a refund, they should take as much as they can, that is $q_{\max}$. Then the buyer makes a payment congruent with the gas purchased to meet the condition of refund taking. However, the buyer may lose money if he/she decides to take $q_{\max}$. In this transaction, the payment

equals $((MB_L + C) - Q) \cdot I$. At the same time, the buyer receives gas which has a value of $q_{\max} \cdot S$.

* If $((MB_L + C) - Q) \cdot I > q_{\max} \cdot S$, that is the buyer loses money through this transaction, then

$$q^* (S, I, Q, C, M, t_{L,J}) = q_{\min}.$$

* If $((MB_L + C) - Q) \cdot I \leq q_{\max} \cdot S$, that is the buyer makes a profit through this transaction, then

$$q^* (S, I, Q, C, M, t_{L,J}) = q_{\max}.$$

– If $(MB_L + C) - Q \geq q_{\max}$, the buyer cannot get a refund, even if the daily maximum quantity of gas has been purchased, and they should therefore take the daily minimum.

$$q^* (S, I, Q, C, M, t_{L,J}) \quad = \quad q_{\min}.$$

Dynamic programming Once we have the terminal optimal decision $q^*_{t_{L,J}} = q^*(S, I, Q, C, M, t_{L,J})$, we can obtain the terminal value of the contract

$$\begin{aligned} V(S, I, Q, C, M, t_{I,J}) =& q^*_{t_{L,J}} \cdot (S - I) \\ &+ \eta \cdot I \cdot \min\left\{Q + q^*_{t_{L,J}} - (MB_L - C), 0\right\} \\ &+ I \cdot \min\left\{M, \max\left\{Q + q^*_{t_{L,J}} - (MB_L + C), 0\right\}\right\}. \end{aligned}$$

Then we work backwards in time. On the last day of each year i (except the final year) the buyer should choose the optimal exercise quantity, the optimal usage of the carry-forward bank and the optimal usage of the make-up bank

$(q^*, c^*, m^*)(S, I, Q, C, M, t_{i,J})$ according to

$$\begin{aligned}
&(q^*, c^*, m^*)(S, I, Q, C, M, t_{i,J}) \\
=&\text{argmax}_{q_{t_{i,J}}, c_i, m_i} \Big\{ q_{t_{i,J}} (S - I) \\
&+ \eta \cdot I \cdot \min \Big\{ Q + q_{t_{i,J}} - (MB_i - c_i), 0 \Big\} \\
&+ I \cdot \min \Big\{ m_i, \max \left\{ Q + q_{t_{i,J}} - (MB_i + C), 0 \right\} \Big\} \\
&+ \mathbb{E} \left[e^{-r\triangle t} V(S_{i+1,1}, I_{i+1,1}, \tilde{Q}_i, C_{i+1}, M_{i+1}, t_{i+1,1}) \mid S_{i,J} = S, I_{i,J} = I \right] \Big\},
\end{aligned}$$

where $\tilde{Q}_i$ is the existing volume of gas which has been taken at the beginning of year i. Once the optimal decisions $(q^*_{t_{i,J}}, c^*_i, m^*_i) = (q^*, c^*, m^*)(S, I, Q, C, M, t_{i,J})$ are determined, we have the contract value

$$\begin{aligned}
&V(S, I, Q, C, M, t_{i,J}) \\
=&q^*_{t_{i,J}}(S - I) + \eta \cdot I \cdot \min \Big\{ Q + q^*_{t_{i,J}} - (MB_i - c^*_i), 0 \Big\} \\
&+ I \cdot \min \Big\{ m^*_i, \max \left\{ Q + q^*_{t_{i,J}} - (MB_i + C), 0 \right\} \Big\} \\
&+ \mathbb{E} \left[e^{-r\triangle t} V(S_{i+1,1}, I_{i+1,1}, \tilde{Q}_i, C_{i+1}, M_{i+1}, t_{i+1,1}) \mid S_{i,J} = S, I_{i,J} = I \right].
\end{aligned}$$

On each day within a gas year, the buyer should choose the optimal quantity $q^*(S, I, Q, C, M, t_{i,j})$ according to

$$\begin{aligned}
&q^*(S, I, Q, C, M, t_{i,j}) \\
=&\text{argmax}_{q_{t_{i,j}}} \Big\{ q_{t_{i,j}}(S - I) + \\
&\mathbb{E} \left[e^{-r\triangle t} V(S_{i,j+1}, I_{i,j+1}, Q + q_{t_{i,j}}, C, M, t_{i,j+1}) \mid S_{i,j} = S, I_{i,j} = I \right] \Big\}.
\end{aligned}$$

Again, once we have the optimal quantity $q^*_{t_{i,j}} = q^*(S, I, Q, C, M, t_{i,j})$, we have the contract value

$$\begin{aligned}
&V(S, I, Q, C, M, t_{i,j}) \\
=&q^*_{t_{i,j}}(S - I) + \\
&\mathbb{E} \left[e^{-r\triangle t} V(S_{i,j+1}, I_{i,j+1}, Q + q^*_{t_{i,j}}, C, M, t_{i,j+1}) \mid S_{i,j} = S, I_{i,j} = I \right].
\end{aligned}$$

TABLE 3.1: Parameters set for Contract **A** and Contract **B**.

Contract **A** with a constant strike price $K = 100$				
$\alpha_S = 5$	$\sigma_S = 0.5$	$r = 0.05$	$q_{\min} = 0$	$q_{\max} = 1$
$MB_1 = 273$	$ACQ_1 = 365$	$\eta = 1$	$F^S(0,t) = 100$, for $0 \leq t \leq L$	
Contract **B** with indexation				
$\alpha_I = 15$	$\sigma_I = 0.2$	$\rho = 0.5$	$F^I(0,t) = 100$, for $0 \leq t \leq L$	

Remark 3.3. Recall that the bang-bang consumption is introduced in Section 2.1. The bang-bang consumption means that the optimal decision of the GSA on each day is either the daily maximum or the daily minimum. In Edoli et al. (2013), the authors show that GSAs with both hard constraints and make-up banks have the bang-bang consumption (similar results can be found in Edoli and Vargiolu (2013)). For GSAs with penalties, make-up banks and carry-forward banks, although no theoretical result available, we observe the bang-bang consumption in our numerical examples.

3.4 Numerical analysis

In this chapter, we evaluate the price of the gas sales agreement on a two-dimensional trinomial tree. We first construct two fundamental trees for both the gas price and the index by assuming $\theta^S(t) = \theta^Z(t) = 0$ in (3.5) and (3.6). Then we shift the nodes on the these two fundamental trees by adding proper drifts in order to be consistent with the observed forward curves. The tree building procedures can be found in Section 2.2.

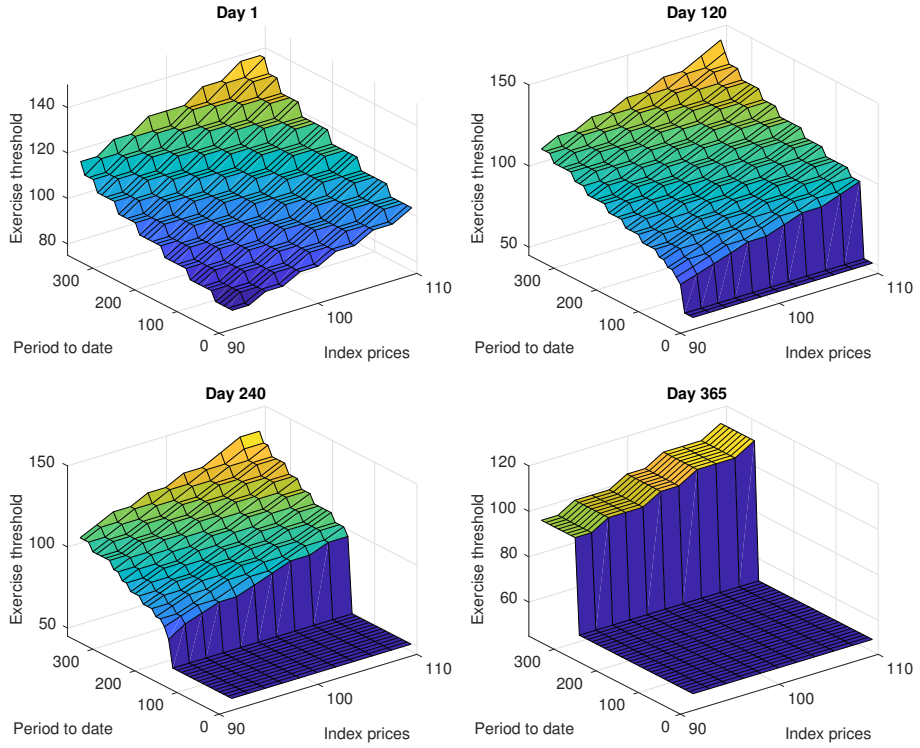
3.4.1 Comparison with the constant strike GSA

In this subsection, we provide two contracts, one with a constant strike price and one with indexation. These contracts last one year with daily exercise opportunities. That is, $L = 1$ and $J = 365$. We show the features of the daily decision surface and the difference between these two contracts. The parameter inputs of those two contracts are shown in Table 3.1.

Contract **A** is a normal one-year contract with daily exercise opportunities and an at-the-money constant strike, $K = 100$. Since we focus on the strike price, we use a flat gas price forward curve to minimize the effect of

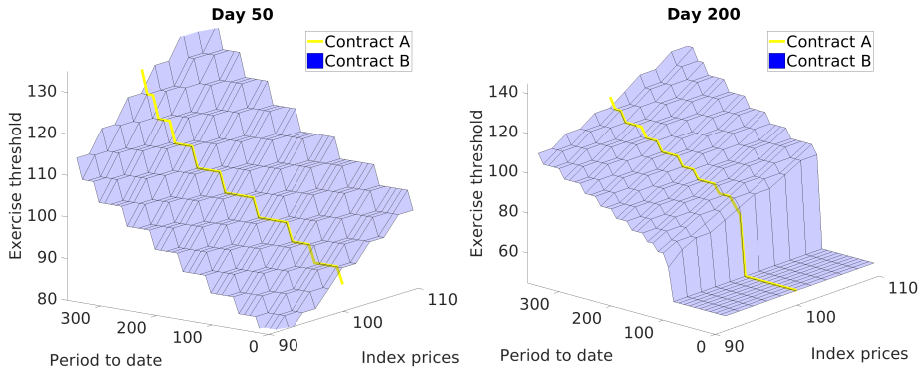
the gas price. In addition, we let the daily take constraints be $q_{\min} = 0$ and $q_{\max} = 1$. Contract **B** uses the same parameter inputs as Contract **A**, except for the constant strike K. Since the index is driven by (3.6), we have those inputs in Table 3.1. For the index process, we choose a relatively large mean-reverting speed and small volatility since the index is determined by the weighted average prices of some other energy products in order to reduce uncertainty. Also, to offer a relatively fair comparison, we let the index forward curve be $F^I(0,t) = K = 100$, for $0 \leq t \leq L$.

FIGURE 3.1: Decision surfaces for Contract **B**



Decision surfaces Since we have observed the bang-bang consumption in our numerical examples, we define the exercise threshold as the gas price equal to or above which the buyer would be better off taking the daily maximum $q_{\max}$. A low exercise threshold means that the buyer is willing to buy

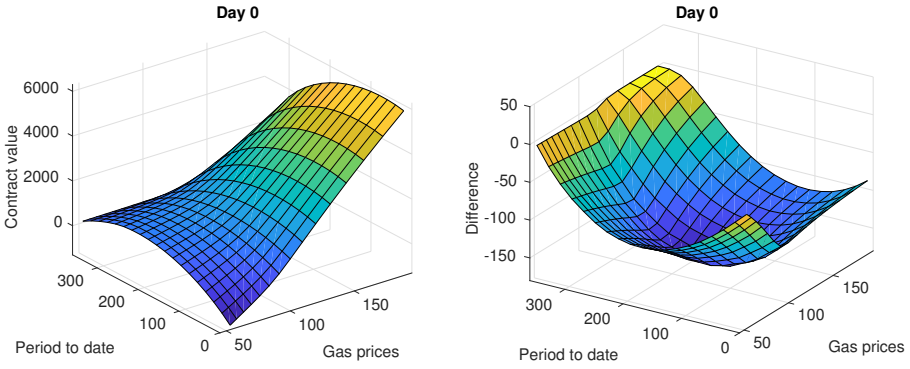
FIGURE 3.2: Comparison of decision surfaces between Contract **A** and Contract **B**



gas under the contract, while a high exercise threshold means the buyer is reluctant to purchase gas. For the buyers, the most important thing is to make the optimal daily exercise decision at the beginning of each gas day, based on the current gas price and the index price. In Figure 3.1 we present some decision surfaces for Contract **B**. Note that the buyer and supplier can specify an existing volume of gas which has been taken before entering a GSA contract. That is why the "Day 1" surface also provides decisions for the period-to-date greater than zero. As we can observe in Figure 3.1, Contract **B** captures almost all the key features of the decision surfaces of a typical GSA with a constant strike (for the analysis of decision surfaces of constant strike GSAs, we refer interested readers to Breslin et al. (2008)). In these plots, the exercise threshold is below the index price when the period to date is low, which would cause an immediate loss. This is due to the minimum bill condition. If the buyers try to avoid any loss at this point then they would have to face more losses in order to meet the minimum bill, or even face penalties, in the future. In addition, as is shown in the "Day 120", "Day 240" and "Day 365" plots, as the time approaches the year end, there will be a quantity of the period to date below which the exercise threshold is the minimum gas price obtained from the tree, regardless of the index price. This is due to the need to meet the minimum bill or reduce the possible penalty. We call this quantity the minimum bill constrained quantity.

Next, we present the difference in decision surfaces between Contract **A**

FIGURE 3.3: Comparison of value surfaces



Note: The left plot shows the value surface of Contract **B**. The right plot shows the difference of values between Contract **A** and Contract **B**.

and Contract **B**. As we can see in Figure 3.2, the exercise threshold of Contract **A** (the yellow line) is above the decision surface of Contract **B** when the period to date is high and beneath the surface when the period to date is low. This means that, as long as the period to date exceeds the minimum bill constrained quantity, the Contract **B** holder would purchase more gas when the period to date is high and purchase less gas when the period to date is low. With a high period to date, the possibility of failing to meet the minimum bill, or facing a penalty, is low. In this scenario, the holder of Contract **B** buys more gas, indicating that the indexation increases the uncertainty of future cashflow. In contrast, with a small period to date, the possibility of failing to meet the minimum bill, or facing a penalty, is high. In this scenario, exercising less gas indicates that the holder of Contract **B** benefits from the indexation.

Contract values The left-hand plot in Figure 3.3 shows the value surface at day 0 of Contract **B**. The right-hand one is obtained by using the value of Contract **B** minus the value of Contract **A** at day 0. As we can see, the value of Contract **A** dominates the value of Contract **B** at each point, which means that the GSA contract holder may not benefit from the indexation. Furthermore, it can be seen that the difference between the values of the two

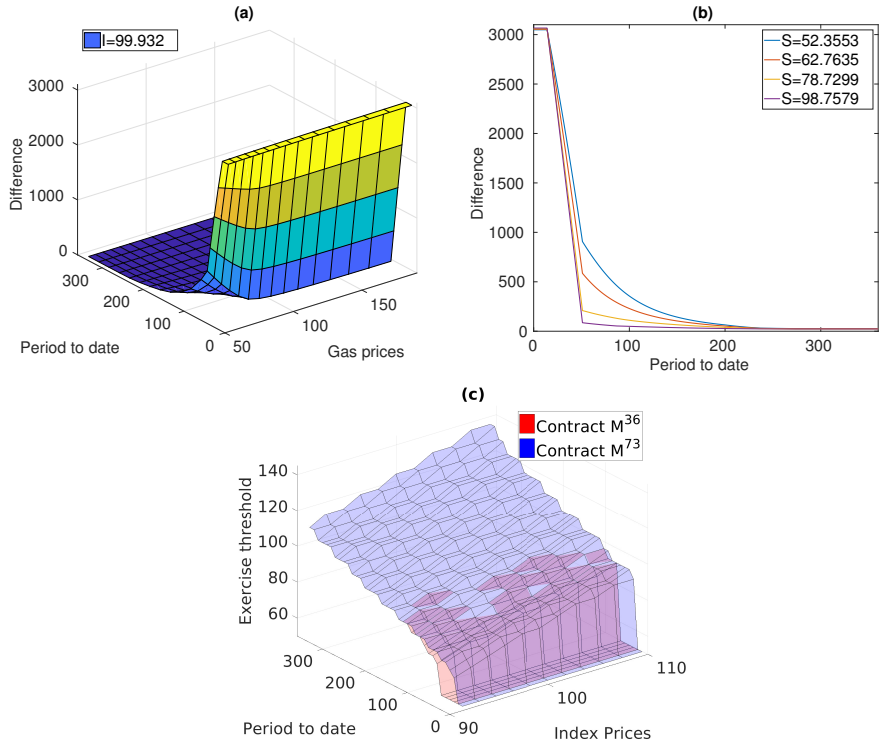
contracts first becomes larger as the period to date increases. This is because, with an initially small period to date, the possible need to meet the minimum bill reduces as the period to date increases. The difference then becomes smaller as the period to date increases. This is due to the fact that the exercise strategies of both contracts become less flexible as the number of possible exercise opportunities decreases. Less flexibility means more similarity, that is, the exercise strategies of both contracts converge to each other as the period to date increases.

3.4.2 The effect of the make-up and carry-forward banks

In GSA contracts, make-up and carry-forward banks offer buyers the opportunity to reduce the potential risk. These rights are constrained, however, by the make-up and carry-forward recovery limits. In this subsection, we investigate the effects of these limits. All contracts appearing in this subsection share the parameters of Contract **B** in Table 3.1. Since the make-up and carry-forward banks appear in multi-year contracts, we let $L = 3$, $J = 365$. In addition, we have the same minimum bill, carry-forward base and annual contract quantity in all three years: $MB_i = 273$, $CB_i = 292$ and $ACQ_i = 365$ for $i = 1, 2, 3$.

The effect of the make-up recovery limit We first present one example where $CRL_i = 0$, $i = 1, 2, 3$. That is, there is no carry-forward clause in the contract. Figure 3.4 shows the difference between Contract $\mathbf{M}^{36}$ with $MRL_i = 36$ and Contract $\mathbf{M}^{73}$ with $MRL_i = 73$, $i = 1, 2, 3$. In (a) and (b), we can see that the value difference increases rapidly if the period to date is less than 51 (i.e. $Q_{t_{1,180}} < 51$). This coincides with our expectations. Intuitively speaking, at the beginning of day 180, once $Q_{t_{1,180}} < 87$, the buyer cannot meet the minimum bill even if he/she takes the maximum daily gas every day for the rest of the current year ($Q_{t_{1,180}} + 1 + (365 - 180) < 273$). It is at this point that the make-up bank starts to work; in other words, the shortfall is added to the make-up bank. Once $Q_{t_{1,180}} < 87 - 36 = 51$, however, the quantity of gas in the make-up bank of Contract $\mathbf{M}^{36}$ will reach the make-up recovery limit, even if the buyer takes the maximum daily gas every day for the rest of the current year. That is, the holder of Contract $\mathbf{M}^{36}$ faces penalties at the year end due to the shortfall, and the extra gas cannot be added to the

FIGURE 3.4: Plots at day 180

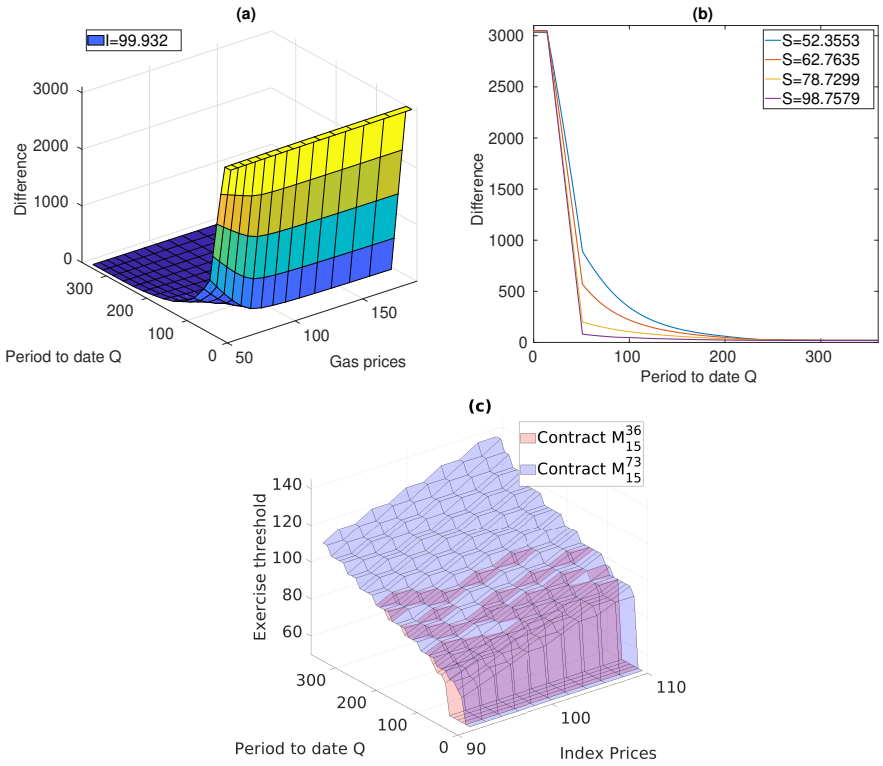


Note: (a) is obtained by using the value of Contract $\mathbf{M}^{73}$ minus the value of Contract $\mathbf{M}^{36}$. By selecting some specified gas prices in (a), we get (b). (c) shows both the decision surfaces for Contracts $\mathbf{M}^{73}$ and $\mathbf{M}^{36}$.

make-up bank and cannot be refunded in the future. At the same time, the shortfall is still being added to the make-up bank of Contract $\mathbf{M}^{73}$, and can be used in future years. The difference then becomes stable when the period to date is less than 14 ($Q_{t_{1,180}} < 51 - (73 - 36) = 14$). At this stage, the gas in the make-up bank of Contract $\mathbf{M}^{73}$ also reaches its recovery limit.

In addition, as we can see in (b) that when the period to date is larger than 51, the difference increases faster in a lower gas price regime than in a higher gas price regime, since the period to date decreases. This means that when the gas price is low, the buyer with more make-up rights has the opportunity to gain more profit by taking less gas in the current year and getting a refund in

FIGURE 3.5: Plots at day 180



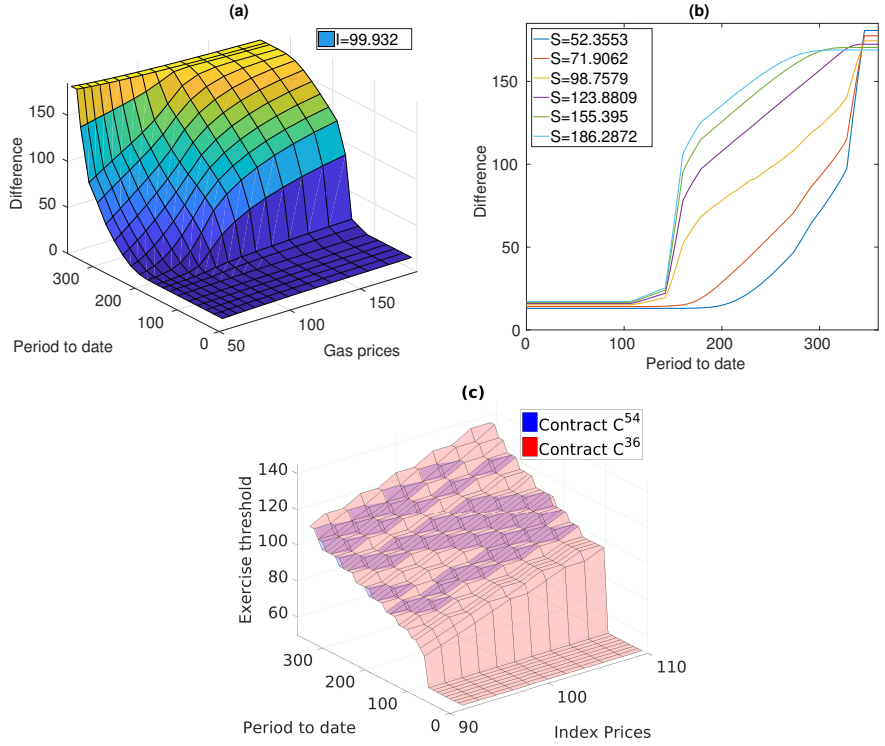
Note: (a) is obtained by using the value of Contract $\mathbf{M}_{15}^{73}$ minus the value of Contract $\mathbf{M}_{15}^{36}$. By selecting some specified gas prices in (a), we get (b). (c) shows both the decision surfaces for Contracts $\mathbf{M}_{15}^{73}$ and $\mathbf{M}_{15}^{36}$.

future years. As we can see in the decision surfaces (Figure 3.4(c)), however, the exercise threshold of Contract $\mathbf{M}^{73}$ is still low, and the buyer with more make-up rights only takes less gas when the period to date is very small. This means that, even with more make-up rights, paying a penalty at year end is still a risky action.

For the sake of completeness, we also provide one example where the carry-forward recovery limit is greater than zero. Figure 3.5 shows the difference between Contract $\mathbf{M}_{15}^{36}$ with $MRL_i = 36$ and Contract $\mathbf{M}_{15}^{73}$ with $MRL_i = 73$, $i = 1, 2, 3$. Both Contract $\mathbf{M}_{15}^{36}$ and Contract $\mathbf{M}_{15}^{73}$ have the carry-forward recovery limit $CRL_i = 15$, $i = 1, 2, 3$. As we can see, Figure 3.4 and

Figure 3.5 hold the same characteristics.

FIGURE 3.6: Plots at day 180



Note: (a) is obtained by using the value of Contract $\mathbf{C}^{54}$ minus the value of Contract $\mathbf{C}^{36}$. By selecting some specified gas prices in (a), we get (b). (c) shows the decision surfaces of both Contracts $\mathbf{C}^{54}$ and $\mathbf{C}^{36}$.

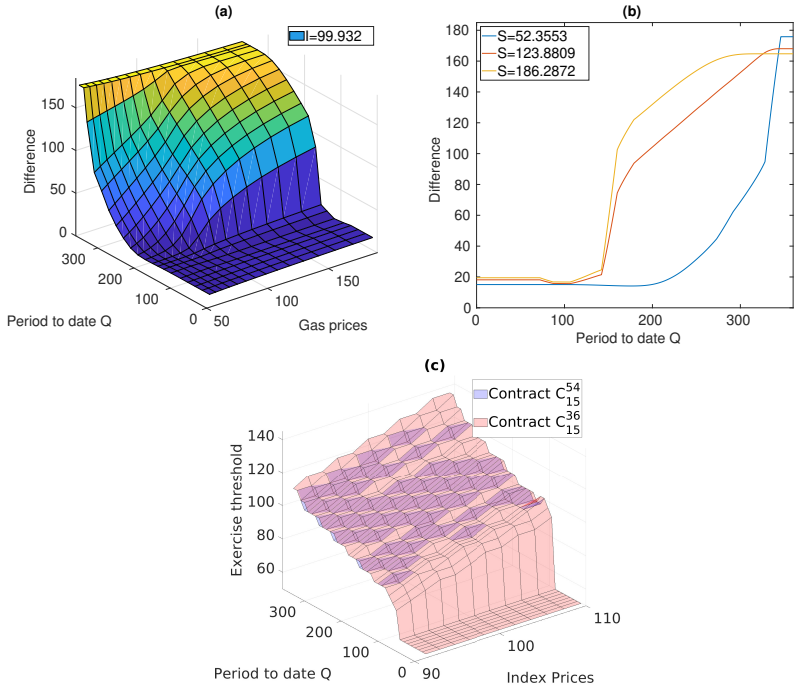
The effect of the carry-forward recovery limit We first present one example where $MRL_i = 0$,$i = 1, 2, 3$. That is, there is no make-up clause in the contract. Figure 3.6 shows the difference between Contract $\mathbf{C}^{54}$ with $CRL_i = 54$ and Contract $\mathbf{C}^{36}$ with $CRL_i = 36$, $i = 1, 2, 3$. In (a) and (b), we can see that the difference starts to increase when the period to date is larger than 106 ($Q_{t_{1,180}} > 106$) in the high gas price regime. This is because the buyer can meet the carry-forward base if he/she takes the maximum daily gas every day for the rest of the current year ($Q_{t_{1,180}} + 1 + (365 - 180) > 292$). At this

stage, both holders of Contract $\mathbf{C}^{54}$ and $\mathbf{C}^{36}$ could have carry-forward gas at the year end, but the difference starts to increase. This indicates that the buyer with more carry-forward rights has the chance to buy more gas when the period to date is relatively low. Once the period to date is larger than 142 ($Q_{t_{1,180}} > 106 + 36 = 142$), the difference starts to increase rapidly in the high gas price regime. This is because, once $Q_{t_{1,180}} > 142$, the quantity of gas in the carry-forward bank of Contract $\mathbf{C}^{36}$ will reach its carry-forward recovery limit if the buyer takes the maximum daily gas every day for the rest of the current year. That is, the extra gas cannot be added to the carry-forward bank of Contract $\mathbf{C}^{36}$ and cannot be used to reduce the minimum bill in future years. At the same time, the extra gas is added to the carry-forward bank of Contract $\mathbf{C}^{54}$, and can be used in future years. The difference then becomes stable when the period to date is larger than 346. At this stage, for both contracts, the gas in the carry-forward bank has already reached the respective recovery limits, regardless of future purchases.

In addition, as we can see in (b), when the carry-forward bank comes into play, the difference is less sensitive with respect to the period to date in the low gas price regime than in the high gas price regime. This is because, with a lower gas price, the possibility of low gas prices in the future is high, and thus the buyer may not buy enough gas to meet the carry-forward base. As we can see in Figure 3.6(c), which shows the decision surfaces for these two contracts, the buyer with more carry-forward rights even takes more gas when the contract is out of the money (i.e. when the exercise threshold is smaller than its corresponding index price). This means that the possible penalty has a big influence on the daily decisions. The buyer intends to do whatever they can to avoid possible penalties. Also, compared with Figure 3.4, it seems that the effect of the carry-forward recovery limit is much less significant than the effect of the make-up recovery limit. This is because the adding of make-up gas would mean an immediate penalty at the current year end, and in future years, the buyer will try to make full use of the make-up gas in order to get refunds. Adding gas to the carry-forward bank, however, usually comes without a loss in the current year, and the buyer may not make full use of the carry-forward gas in the future.

For the sake of completeness, we also provide one example where the make-up recovery limit is greater than zero. Figure 3.7 shows the difference

FIGURE 3.7: Plots at day 180



Note: (a) is obtained by using the value of Contract $\mathbf{C}_{15}^{54}$ minus the value of Contract $\mathbf{C}_{15}^{36}$. By selecting some specified gas prices in (a), we get (b). (c) shows the decision surfaces of both Contracts $\mathbf{C}_{15}^{54}$ and $\mathbf{C}_{15}^{36}$.

between Contract $\mathbf{C}_{15}^{36}$ with $CRL_i = 36$ and Contract $\mathbf{C}_{15}^{54}$ with $CRL_i = 54$, $i = 1, 2, 3$. Both Contract $\mathbf{C}_{15}^{36}$ and Contract $\mathbf{C}_{15}^{54}$ have the make-up recovery limit $MRL_i = 15$, $i = 1, 2, 3$. Generally speaking, Figure 3.6 and Figure 3.7 share the same characteristics. However, we find that Figure 3.7(a) behaves differently from Figure 3.6(a) when the gas price is high. This phenomena becomes clear in Figure 3.7(b). In Figure 3.7(b), the difference increases as the period to date decreases from $Q_{t_{1,180}} = 87$ to $Q_{t_{1,180}} = 72$ when the gas price is high. This is because that, once $Q_{t_{1,180}} < 87$, the buyer cannot meet the minimum bill even if he/she takes the maximum daily gas every day for the rest of the current year ($Q_{t_{1,180}}+1+(365-180) < 273$). It is at this point that the shortfall is added to the make-up banks of both Contract $\mathbf{C}_{15}^{36}$ and Contract

$\alpha_S = 2$	$\sigma_S = 0.5$	$\alpha_I = 3$	$\sigma_I = 0.4$	$r = 0.05$	$\rho = 0.5$
$q_{\min} = 0$	$q_{\max} = 1$	$MB = 39$	$CB = 42$	$ACQ = 52$	$\eta = 1$

TABLE 3.2: Parameter values.

$\mathbf{C}_{15}^{54}$. The difference starts to increase as the period to date decreases indicates that the carry-forward bank slightly amplifies the functionality of the make-up bank. Once $Q_{t_{1,180}} < 87 - 15 = 72$, the gas in the make-up banks of both Contract $\mathbf{C}_{15}^{36}$ and Contract $\mathbf{C}_{15}^{54}$ will reach the make-up recovery limits. This is why the difference stops increasing when the period to date is less than 72.

3.4.3 The effect of various parameters

In this subsection, we provide a detailed analysis of how various parameters affect the contract values.

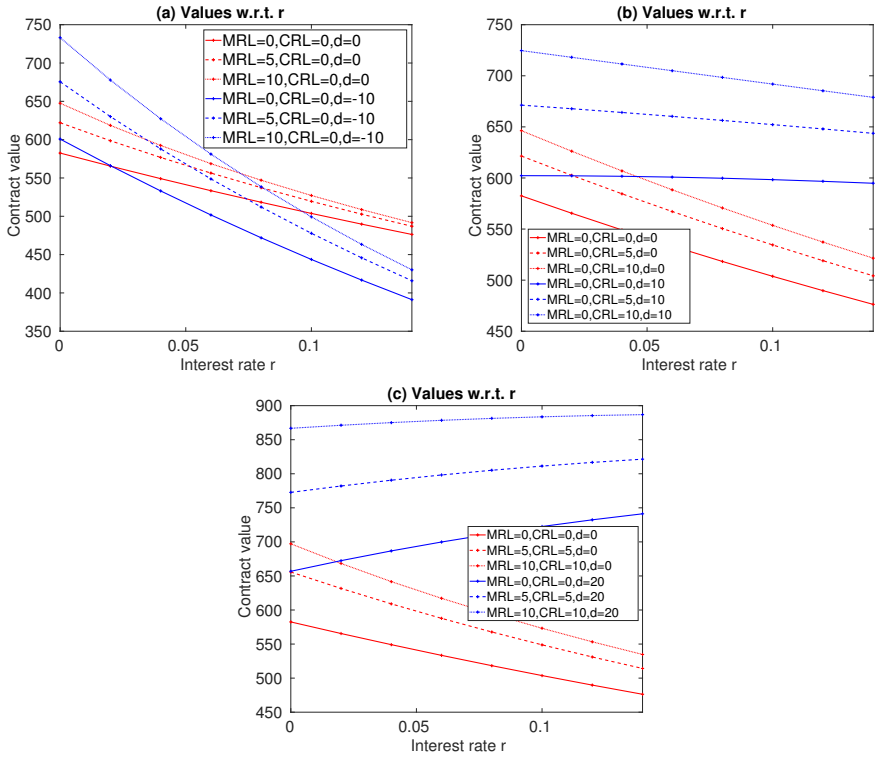
Parameter settings We choose a three-year GSA contract with weekly exercise opportunities. That is $L = 3$, $J = 52$. When the parameters are not variable, we use those values in Table 3.2. By the nature of make-up and carry-forward banks, for such a three-year example, the make-up bank will play a large role if the contract is out-of-the-money in the first year and at-the-money or in-the-money in the future years. Similarly, the carry-forward bank would contribute more if the contract is in-the-money in the first year and at-the-money or out-of-the-money in the future years. Thus one cannot have a fair analysis without taking the forward curves into consideration, since the forward curves can control the price movement in some sense. Although we can adjust either the gas forward curve or the index forward curve, since the index is obtained by the average of other energy substitutes, and is therefore more stable than the gas price, we choose the gas forward curve so as to model a changeable situation. We use a flat index forward curve, that is $F^I(0,t) = 100$, for $0 \le t \le L$. Let the gas forward curve be given as follows:

$$F^S(0,t) = \begin{cases} 100 + d, & 0 \le t \le 1, \text{ Year } 1, \\ 100, & 1 < t \le 2, \text{ Year } 2, \\ 100 - d, & 2 < t \le 3, \text{ Year } 3, \end{cases}$$

where $d \in (-100, 100)$ is a constant. We call d the decoupling level. When $d < 0$, then the contract is likely to be out-of-the money in the first year, at-the-money in the second year and in-the-money in the third year. In this situation, we say the gas forward curve is negatively decoupled. Curve I and Tree I in Figure 2.4 present an example of a negatively decoupled gas forward curve and its corresponding trinomial tree. When $d = 0$, then the contract is likely to be at-the-money in all three years. In this situation, we say that the gas forward curve is not decoupled. Curve II and Tree II in Figure 2.4 present an example of a gas forward curve that is not decoupled, along with its corresponding trinomial tree. When $d > 0$, then the contract is likely to be in-the-money in the first year, at-the-money in the second year and out-of-the-money in the third year. In this situation, we say that the gas forward curve is positively decoupled. Curve III and Tree III in Figure 2.4 present an example of a positively decoupled gas forward curve, along with its corresponding trinomial tree.

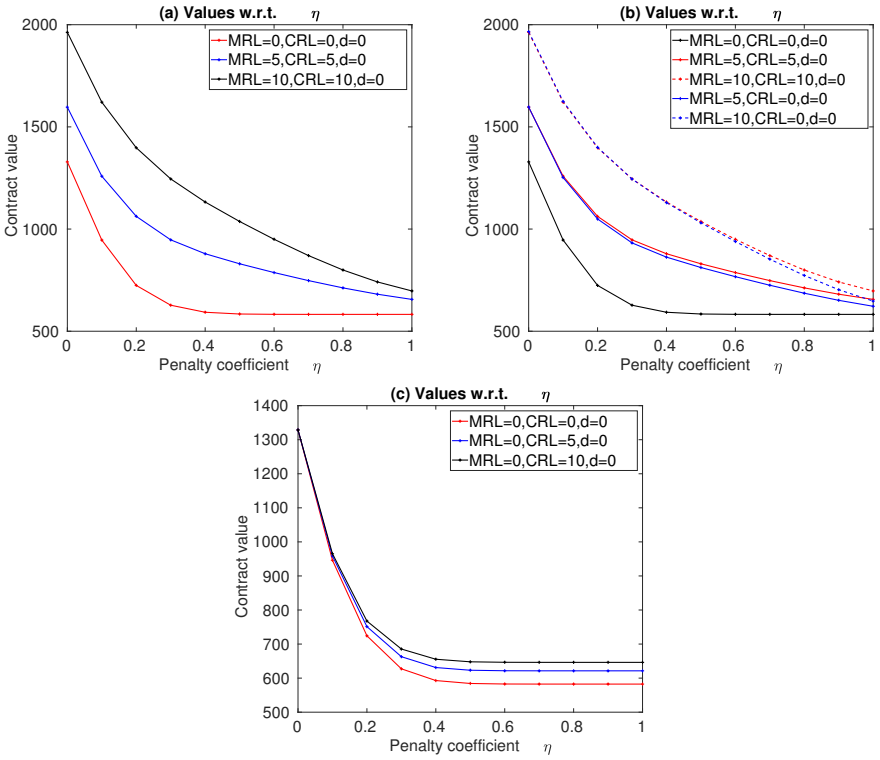
The risk-free rate By the nature of the make-up bank, the addition of make-up gas in some years comes with penalties at the year end. The make-up gas is refunded in future years, which means that there is a time gap between the penalty and the refund. Thus the benefit of the make-up bank is affected by the risk-free rate. Intuitively speaking, the higher the risk-free rate is, the less contribution the make-up bank makes. Recall that the make-up and carry-forward rights are measured by the make-up and carry-forward recovery limits. The larger these limits are, the more make-up and carry-forward rights the holder has. Furthermore, a contract that has larger make-up and carry-forward recovery limits should also have a larger contract value. Figure 3.8(a) evidences our intuitions. The benefit of the greater make-up rights peaks when the risk-free rate is 0, and reduces as the risk-free rate increases. For the carry-forward bank, although there is no penalty when the holder gains carry-forward gas, this gas carried forward could be used to reduce the minimum bill in the future years, and when the carry-forward gas is withdrawn, the possible loss is reduced. There is also a time gap between the addition and withdrawal of carry-forward gas. Thus, the benefit of the carry-forward bank should also be reduced by a higher risk-free rate. Figure 3.8(b)

FIGURE 3.8: The contract value w.r.t. r



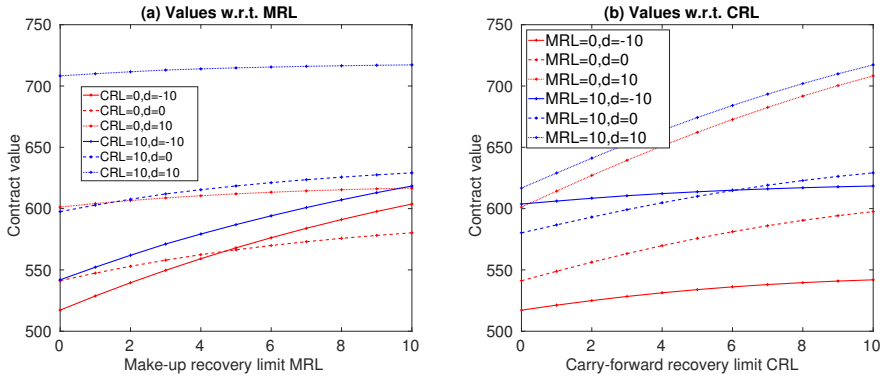
shows these intuitions. As we can see, the benefit of more carry-forward becomes less significant when the risk-free rate increases. Also, compared with the positively decoupled gas forward curve ($d = 10$), these phenomena are less significant when the forward curve is not decoupled ($d = 0$). This is because the holder may have fewer opportunities to withdraw carry-forward gas with $d = 0$. As we can see in Figure 3.8(c), the benefits of both make-up and carry-forward banks are reduced by a higher risk-free rate. In addition, when the forward curve is highly positively decoupled ($d = 20$), the contract value increases as the risk-free rate increases. This is due to the fact that the contract is highly in-the-money in the first year and highly out-of-the-money in the third year. A higher risk-free rate, therefore, amplifies the benefit of the in-the-money period.

FIGURE 3.9: The contract value w.r.t. η



The penalty coefficient To minimize the time value of money, we let the risk-free rate be $r = 0$. The adding of make-up gas in some years comes with penalties at the year end, which means the penalty coefficient should have a significant effect on the make-up clause. A smaller penalty coefficient leads to smaller possible penalties, and thus to buyers potentially accruing more benefit from their make-up rights. Intuitively speaking, the greatest potential benefit arising from make-up rights is probably when η equals zero, since at that level buyers can store make-up gas without paying penalties. As we can see in plot Figure 3.9(a), however, the benefit arising from having more make-up rights is at its greatest when the penalty coefficient is around 0.3. Furthermore, as we can see in Figure 3.9(b) and (c), the penalty coefficient also affects the carry-forward bank. When $\eta = 0$, the contract values are the same regardless of how many carry-forward rights the holder has (Figure 3.9(c)). This is because the carry-forward bank loses its function when $\eta = 0$;

FIGURE 3.10: The contract value w.r.t. MRL and CRL.

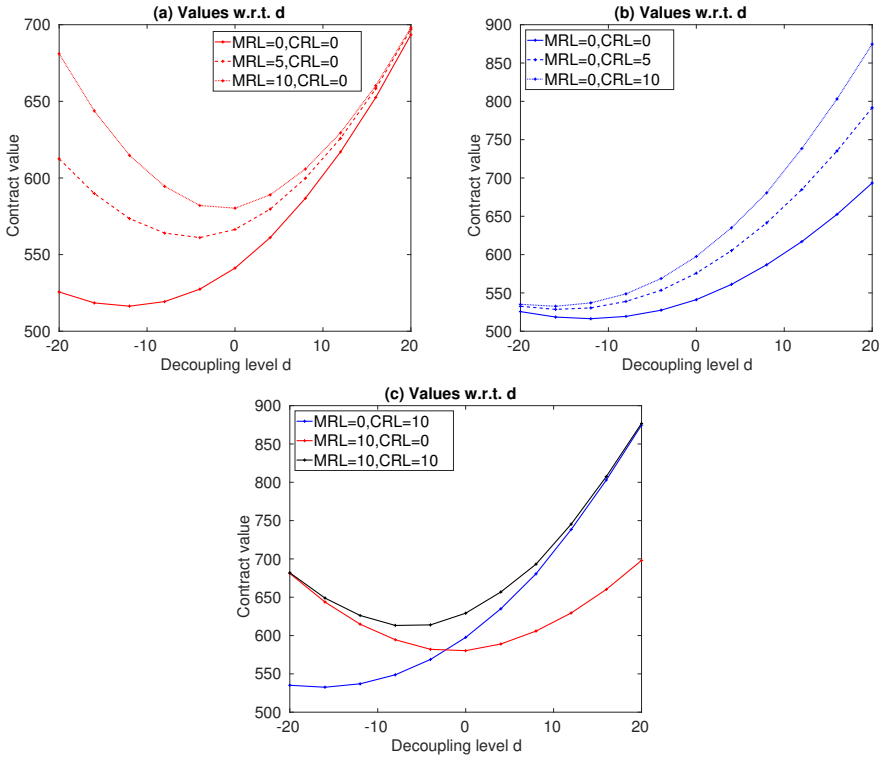


there is no need to reduce the minimum bill by using the carry-forward gas if there is no penalty. The gap between these lines becomes stable when the penalty coefficient is larger. This is because, at this point, the penalty is large enough to maximize the carry-forward bank's functionality. It also suggests that, without a make-up bank, when η is large enough, the buyer will do whatever they can to avoid possible penalties.

The make-up and carry-forward recovery limits Figure 3.10(a) shows that the contract value increases as the make-up recovery limit increases. It also evidences that the increasing speed is faster with a negatively decoupled forward curve than a positively decoupled curve, or a curve that is not decoupled. Figure 3.10(b) shows that the contract value increases as the carry-forward recovery limit increases. It also shows that the increasing speed is faster with a positively decoupled forward curve.

The decoupling level As can be seen in the previous figures, with make-up and carry-forward clauses, the decoupling level has a huge impact on the contract. Figure 3.11(a) and (b) shows that the benefit of more make-up rights increases as the decoupling level d decreases, while the benefit of more carry-forward increases as the decoupling level increases. The more the forward curve is negatively decoupled, the more make-up gas the holder can store in the first year and withdraw in the third year when the contract is highly

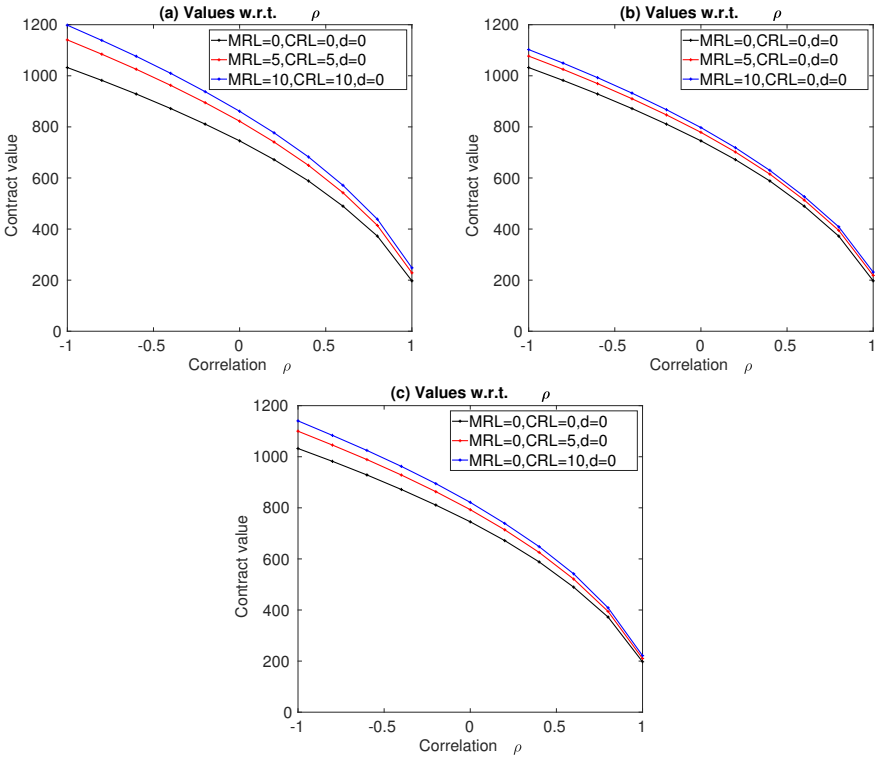
FIGURE 3.11: The contract value w.r.t. d



in-the-money. When the forward curve is highly positively decoupled, meanwhile, the holder can buy more gas and gain more carry-forward gas, since when the contract is highly out-of-the-money in the third year, the holder will have more carry-forward gas to use to reduce the minimum bill. Figure 3.11(c) also evidences that, when the forward curves are not decoupled, the carry-forward bank contributes more to the contract value than the make-up bank.

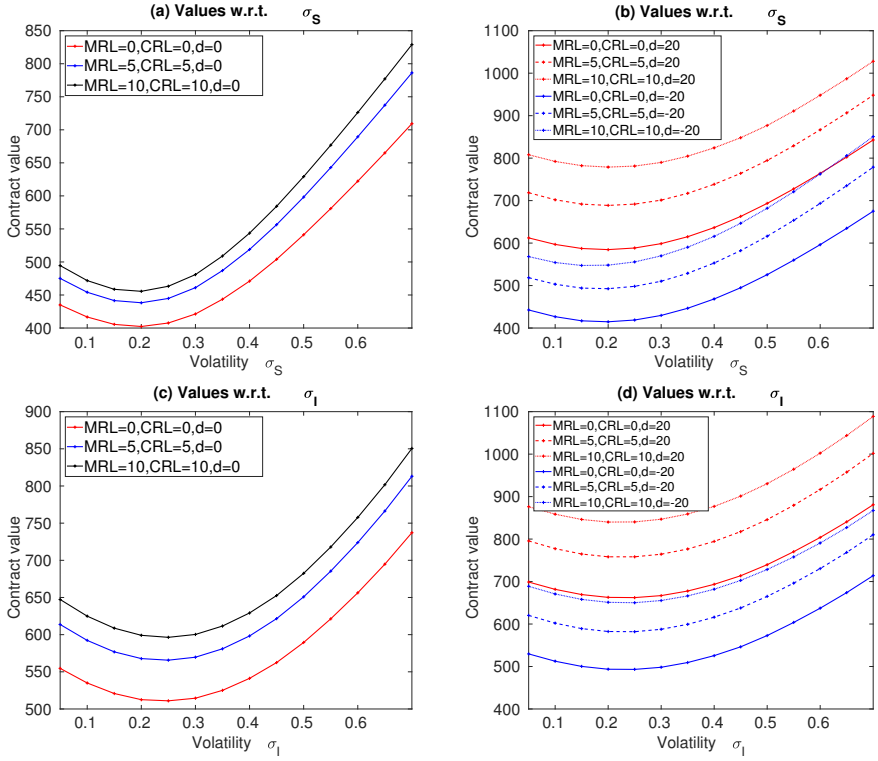
The correlation The effect of the correlation ρ is presented in Figure 3.12. It can be observed that the contract value decreases when the correlation increases. This means that the buyer can benefit more when the gas price and the index are negatively correlated. Figure 3.12(b) and (c) also show that a positive correlation weakens the function of both the make-up and carry-forward banks.

FIGURE 3.12: The contract value w.r.t. ρ



The volatility In respect to the effect of for the volatility of both the gas price and the index, Figure 3.13 shows that the contract value decreases as the volatility increases in a lower volatility regime, but then increases in a higher volatility regime. In addition, the contract value peaks at the largest value of the volatility. This is because the buyer could implement a more flexible trading strategy when the gas price or the index price is fluctuating more. Figure 3.13(a) also shows that, when the forward curve is less decoupled (the absolute value of d is small) the holder with more make-up and carry-forward rights would benefit more from higher gas price volatility, while Figure 3.13(c) shows that one would not appreciate this feature when it comes to the volatility of the index. As we can see in Figure 3.13(b), when the forward curve is highly decoupled (the absolute value of d is large), the holder with more make-up and carry-forward rights would not gain any significant benefit from higher volatility. This is because the highly decoupled

FIGURE 3.13: The contract value w.r.t. σ_S and σ_I



curve has already made sure the make-up or carry-forward bank works to its full capacity.

The mean-reverting speed Figure 3.14 shows that the contract value is decreasing in α_S and increasing in α_I. This also shows that the speed of reversion to the mean for both gas price and index does not have a significant impact on the benefit of having more make-up and carry-forward rights.

3.4.4 How the indexation affects the decisions

In this section, we show how the indexation affects the decisions on weekly exercise, carry-forward bank and make-up bank, and also the influences on the volume taken. We use a six-year GSA contract with weekly exercise opportunities ($L = 6$, $J = 52$) as an example. The forward curves for both gas

FIGURE 3.14: The contract value w.r.t. α_S and α_I

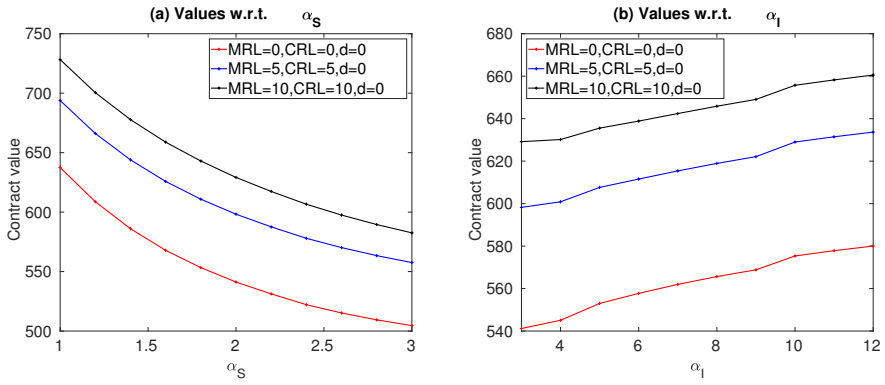
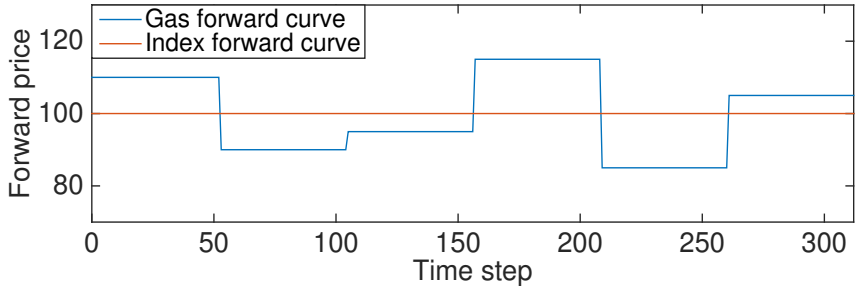


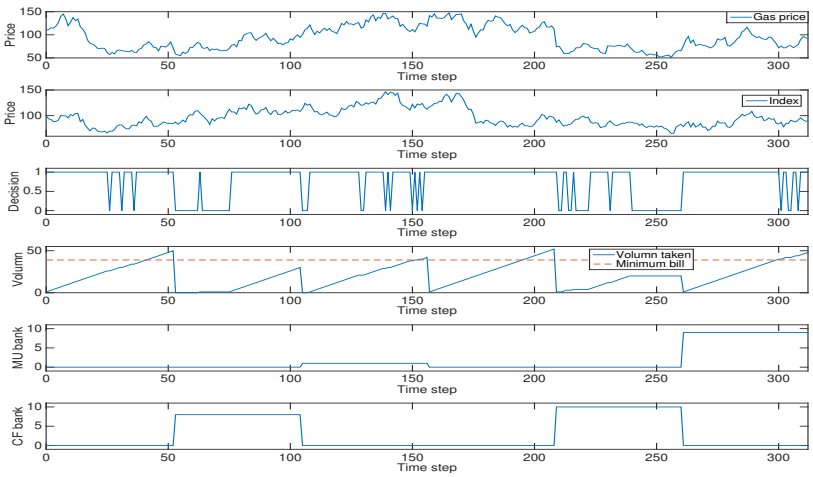
FIGURE 3.15: Forward curves



price and index are given in Figure 3.15. Using the parameters in Table 3.2, we let $MRL_i = 10$ and $CRL_i = 10$ for $i = 1, \ldots, 6$. We simulate a path of gas spot prices and two paths of index at the same time. Then we make decisions based on the optimal decision surface we calculated by using the dynamic programming technique in Section 3.3. Figures 3.16 and 3.17 demonstrate how decisions change when the realizations of the index are different.

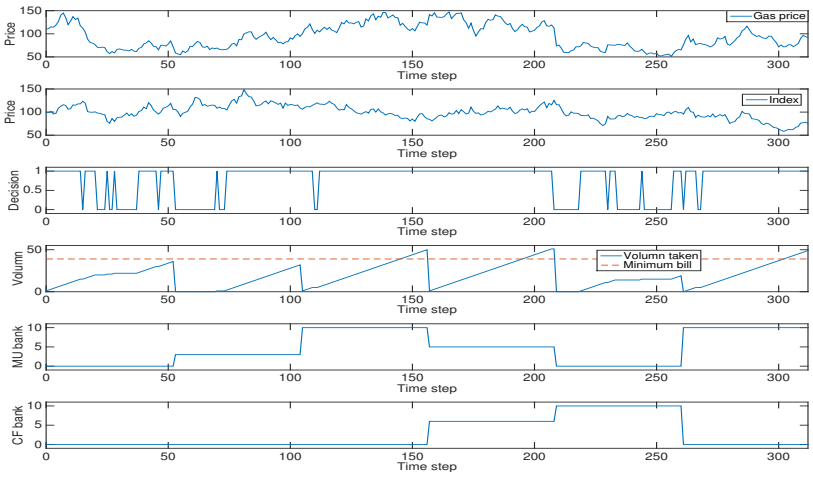
In the first year of this example the gas forward curve is above the index forward curve (i.e. in-the-money). Assuming that the spot price and the index both follow their forward curves (we call this the intrinsic strategy), the optimal strategy would then be to take the maximum possible ($Q_{T_1} = 52$) and create ten units of carry-forward ($C_2 = 52 - 42 = 10$). In Figure 3.16, however, the simulated take is 50 ($Q_{T_1} = 50$), which creates eight units of carry-forward

FIGURE 3.16: One realization



Note: One realization of the index and the corresponding optimal takes, volume taken, and the evolution of both make-up and carry-forward banks.

FIGURE 3.17: Another realization



Note: Another realization of the index and the corresponding optimal takes, volume taken, and the evolution of both make-up and carry-forward banks.

gas ($C_2 = Q_{T_1} - 42 = 8$). While in Figure 3.17, the simulated take is 36 ($Q_{T_1} = 36$), which creates three units of make-up gas ($M_2 = 39 - Q_{T_1} = 3$).

In the second year, the gas forward curve is below the index curve (i.e. out-of-the-money) and so the optimal intrinsic strategy would be to use all the carry-forward gas ($c_2 = 10$) so as to reduce the minimum bill to 29 ($39 - c_2 = 29$) and then take 19 ($Q_{T_2} = 19$) to create a make-up bank of ten units ($M_3 = 39 - c_2 - Q_{T_2} = 10$, the maximum that can be recovered in the coming year). In Figure 3.16, the simulated strategy uses all the carry-forward gas ($c_2 = 8$) to reduce the minimum bill to 31 ($39 - c_2 = 31$) and taking 30 ($Q_{T_2} = 30$), which also creates a make-up bank of one unit ($M_3 = 31 - Q_{T_2} = 1$). In the case of Figure 3.17, the simulated take is 32 ($Q_{T_2} = 32$), increasing the make-up bank to ten units ($M_3 = M_2 + (39 - 32) = 10$).

The third year is also out-of-the-money, so the intrinsic strategy is to take the minimum bill plus the amount of gas in the make-up bank ($m_3 = 10$) that will be refunded, giving a take of 49 ($Q_{T_3} = MB + m_3 = 49$). In Figure 3.16, the simulated strategy is to take 42 ($Q_{T_3} = 42$) and use one unit of make-up gas ($m_3 = 1$), which makes the make-up bank next year to be 0 ($M_4 = M_3 - m_3 = 0$). The simulated take in Figure 3.17 is $Q_{T_3} = 50$ but we only use five units of gas ($m_3 = 5$) in the make-up bank, thus reducing the make-up bank to 5 ($M_4 = M_3 - m_3 = 5$). We also create six units of carry-forward gas ($C_4 = Q_{T_3} - \max\{39 + m_3, 42\} = 6$).

In the fourth year the contract is in-the-money, so the optimal intrinsic strategy is to take the maximum quantity of gas ($Q_{T_4} = 52$) and create another ten units of carry-forward ($C_5 = 10$). The simulated strategy in Figure 3.16 is to take 52 ($Q_{T_4} = 52$), which also creates ten units of carry-forward gas ($Q_{T_4} - 42 = 10$). The simulated strategy in Figure 3.17 is also to take 51 ($Q_{T_4} = 51$) and use all the make-up gas ($m_4 = M_4 = 5$), and thus create a carry-forward bank of ten units for the next year ($C_5 = \max\{C_4 + Q_{T_4} - \max\{39 + m_4, 42\}, CRL_5\} = 10$).

In the fifth year the contract is out-of-the-money, so the intrinsic strategy would be the same as that seen in year two. The simulated strategy in Figure 3.16 is to use all the carry-forward gas ($c_5 = 10$) to reduce the minimum bill to 39-10=29 and take 20 units of gas ($Q_{T_5} = 20$) which makes the make-up bank for the next year to be nine units ($M_6 = 29 - Q_{T_5} = 9$). The simulated strategy in Figure 3.17, meanwhile, is to use ten units of the carry-forward

gas ($c_5 = 10$) to reduce the minimum bill to 39-10=29 and take 19 units of gas ($Q_{T_5} = 19$), which makes the make-up bank next year to be ten units ($M_6 = 29 - Q_{T_5} = 10$).

In the final year the contract is in-the-money, so the intrinsic strategy would be to take the maximum possible, with ten units being taken from the make-up bank. The simulated strategy in Figure 3.16 is to take 52 ($Q_{T_6} = 52$) and use all the make-up gas, which leaves no gas left in either the make-up or carry-forward banks. The simulated strategy in Figure 3.17 is to take 50 ($Q_{T_6} = 50$), which also leaves no gas in either the make-up or carry-forward banks.

3.5 Conclusion

In this chapter, we have proposed a two-dimensional trinomial tree framework for pricing multiple year GSAs with make-up, carry-forward and indexation, given knowledge of forward price dynamics of both gas and index. GSAs are complicated to evaluate both because the buyers can exercise their rights in a daily manner while making decisions on the make-up bank and carry-forward bank on a yearly basis, and because the strike prices are able to move stochastically. Hence, in the evaluation, we need to keep track of multiple variables on a daily basis over a number of years. These complexities require efficient numerical procedures to value these contracts, and herein lies the main contribution of this chapter.

With the help of a two-dimensional trinomial tree, we are able efficiently to evaluate the prices of the contracts so as to find both the optimal daily decisions and the optimal yearly use of both the make-up bank and carry-forward bank. We also demonstrate various features of this complex contract with the help of a number of numerical studies. For example, a buyer with more make-up rights only intends to take less gas when the period to date is very small; whereas a buyer with more carry-forward rights intends to take more gas even when the contract is out-of-the-money. A high interest rate weakens the functionality of both the make-up and carry-forward banks, while a large penalty coefficient increases the value of the carry-forward bank and decreases the value of the make-up bank. The values of both make-up and carry-forward banks are most affected by the decoupling level; and when

the forward curves are not decoupled, the carry-forward bank contributes more to the contract value than the make-up bank.

It should be noted that the definition and properties of the index have been simplified in this chapter. The index in the current month in a real contract is determined by the weighted average of some other energy prices in the previous month, which links the valuation to the moving average problem. These non-Markovian and non-continuous properties make the evaluation of such a real contract much harder than the scenario envisaged in the current chapter, we investigate this problem in Chapter 4, Chapter 5 and Chapter 6.

Chapter 4

Evaluation of gas sales agreements with indexation

4.1 Introduction

When the strike price of a GSA is a constant, the valuation of the GSA is a classic dynamic programming problem. When the strike price is set based on the indexation principle, however, it is called the index. In each month, the value of the index is determined by the weighted average price of some energy products in the previous month (see Asche, Osmundsen and Tveterås (2002) for details). This feature links the evaluation of the GSAs to the moving average problem. Works on general moving average pricing can be found in Dai, Li and Zhang (2010) (a binomial tree approach), Bernhart, Tankov and Warin (2011) (a simulation-based approach) and Xu, Hong and Qin (2013) (a willow tree approach).

There is very limited research in the literature that discusses GSAs with indexation. Edoli et al. (2013) and Dong and Kang (2018) implement GSA evaluation using a two-dimensional tree by assuming that the index process is Markovian. Edoli (2013) relaxes this assumption by assuming that the moving average process is Markovian and implements it using the finite difference method. In Bernhart (2011), the author uses the approximation method in Bernhart, Tankov and Warin (2011) to model the index and implements it using the least-squares Monte Carlo approach. To the author's knowledge, no efficient method has been derived without making those assumptions. In this chapter, inspired by the lattice algorithm used to value the Asian options

(see Dai, Li and Zhang (2010) and Klassen (2001)), we propose a new numerical method to fill the gap.

The chapter is organized as follows: mean-reverting models for the gas price and the oil price are suggested in Section 4.2. In Section 4.3, we formulate the detail of the index price based on the indexation principle in the real GSAs. We generalize and formulate the key features of a GSA with indexation in Section 4.4. Section 4.5 formulates the gas price, the oil price, the index and the running average variable using a two-dimensional tree and generates the grids for both the running average variable and the index. A detailed numerical method for pricing GSAs with indexation is proposed in Section 4.6. Section 4.7 gives a simple example of what the generated grid looks like. Section 4.8 models GSAs with indexation in continuous time and derives the HJB equation which drives the value of a GSA. Section 4.9 proves that the algorithm built in Section 4.6 has the first order consistency to the HJB equation in Section 4.8.

4.2 The pricing framework

Denote the spot prices of gas and oil at time t by $S(t)$ and $Z(t)$, respectively. This chapter assumes that the log prices of the gas and the oil follow the mean-reverting processes

$$d\ln S(t) = \Big[\theta^S(t) - \alpha_S \ln S(t)\Big]dt + \sigma_S dB^S(t), \tag{4.1}$$

$$d\ln Z(t) = \Big[\theta^Z(t) - \alpha_Z \ln Z(t)\Big]dt + \sigma_Z dB^Z(t), \tag{4.2}$$

where $B^S(t)$ and $B^Z(t)$ are standard Brownian motions with correlation ρ, and α_S, σ_S, α_Z and σ_Z are constants. Let $F^S(t,T)$ and $F^Z(t,T)$ be the forward prices of gas and oil at time $t \in [0,T]$ with maturity T, respectively. $\theta^S(t)$ and $\theta^Z(t)$ are functions of time chosen to provide an exact fit to the observed forward curves $F^S(0,t)$ and $F^Z(0,t)$. These processes are standard models when it comes to the commodity price in the energy market (see Edoli et al. (2013) and Hull (2011)).

4.3 The indexation principle

We have the time horizon $[0, T]$ equally spaced into L periods. Let δ be the interval of each period, $\delta = \frac{T}{L}$. Denote the beginning of the l-th period by T_l, $l = 0, 1, \ldots, L-1$, and $T_L = T$. That is,

$$\begin{aligned} 0 = T_0 < T_1 = \delta < T_2 = 2\delta < \ldots \\ < T_l = l\delta < \ldots < T_{L-1} = (L-1)d < T_L = L\delta = T. \end{aligned}$$

Let $I(t)$ be the index at time $t \in [0, T]$. For $t \in [T_0, T_1]$, the index is a fixed constant K which is specified in the contract. For $t \in (T_l, T_{l+1}]$, $l = 1, 2, \ldots, L-1$, the index is the average value of the crude oil price Z in the previous period $(T_{l-1}, T_l]$. That is,

$$I(t) = \begin{cases} K & \text{for} \quad t \in [T_0, T_1], \\ \frac{1}{\delta} \int_{T_{l-1}}^{T_l} Z(u)du & \text{for} \quad t \in (T_l, T_{l+1}],\ l = 1, \ldots, L-1. \end{cases} \tag{4.3}$$

In the period $(T_l, T_{l+1}]$, $l = 0, 1, \ldots, L-2$, since the average value of the oil prices is the index value for the next period, we need another variable, M, to track how the average moves. Thus, we introduce the running average variable M, which is given by

$$M(t) = \frac{1}{t - T_l} \int_{T_l}^{t} Z(u)du. \tag{4.4}$$

where $t \in (T_l, T_{l+1}]$, and $l = 0, 1, \ldots, L-2$, since we do not need to track the running average in the last month. At this point, we can see that the running average variable M is defined on $(0, T_{L-1}]$. In addition, note that M is piecewise continuous on the intervals $(T_l, T_{l+1}]$ for $l = 0, 1, \ldots, L-2$ with a jump at time T_l for $l = 1, \ldots, L-2$. Also, it is straightforward to observe that, for $t \in (T_l, T_{l+1}]$, $l = 1, 2, \ldots, L-1$,

$$\begin{aligned} I(t) &= M(T_l) \\ &= \frac{1}{\delta} \int_{T_{l-1}}^{T_l} Z(u)du, \end{aligned} \tag{4.5}$$

Furthermore, (4.4) gives

$$dM(t) = \frac{Z(t) - M(t)}{t - T_l} dt,$$

where $t \in (T_l, T_{l+1}]$, $l = 0, 1, \ldots, L-2$.

4.4 Gas sales agreements with indexation

We have introduced the one year GSA contract in Section 2.1. Since we have to adopt new notations and build a new framework, we introduce the one year GSA again in this section.

- T is the terminal date of the contract. Since we focus on the one-year GSA, we let $T = 1$. L corresponds to the number of months and δ coincides with the length of each month. That is, $T = L \cdot \delta$. In the rest of this chapter, we call the first month "Month 0", the second month "Month 1", ..., the last month "Month $L-1$".

- Let the period $[T_l, T_{l+1}]$, $l = 0, \ldots, L-1$, be equally spaced into D pieces with interval $\Delta t = \frac{\delta}{D}$, where $T_l = l \cdot \delta$ and D is a positive integer. Then the time horizon $[0, T]$ is equally spaced into $L \cdot D$ pieces. Let $N = L \cdot D$ and $t_n = n \cdot \Delta t$, where $n = 0, 1, \ldots, N$, we have

$$\begin{aligned} T_l = t_{lD} < T_l + \Delta t = t_{lD+1} < T_l + 2\Delta t = t_{lD+2} < \ldots \\ < T_l + (D-1)\Delta t = t_{lD+(D-1)\Delta t} < T_l + D\Delta t = t_{(l+1)D} = T_{l+1}. \end{aligned}$$

- We assume that the holder of a GSA has exactly one exercise right at each time t_n for $n = 1, \ldots, N$, which amounts to D rights for a whole month $(T_l, T_{l+1}]$ and N rights for a whole year $[0, T]$. Typical GSAs can usually be exercised daily (we assume that one year contains 360 days, that is $L = 12$, $D = 30$ and $N = 360$).

- Let q_{t_n} be the amount of gas taken (exercise decision) at time t_n, which is constrained by the minimum daily quantity $q_{\min}$ and the maximum daily quantity $q_{\max}$, that is $q_{\min} \leq q_{t_n} \leq q_{\max}$. Such an admissible policy is denoted by $\mathbf{q} = \{q_{t_n}\}$. Let Q_{t_n} be the cumulative amount of gas taken

before time t_n (also known as the period to date) which is given by

$$Q_{t_n} = \sum_{k=1}^{n-1} q_{t_k} \text{ for } n = 2, \ldots, N,$$

and $Q_{t_n} = 0$ for $n = 0, 1$. In addition, we let Q_T be the total amount of gas taken. That is, $Q_T = Q_{t_N} + q_{t_N}$.

- $S(t_n)$, $Z(t_n)$, $M(t_n)$ and $I(t_n)$ are the gas price, oil price, the value of the running average and the index at time t_n, respectively. Then, upon taking the volume q_{t_n}, the payoff from the buyer's point of view at time $t_n = t_{lD+d} \in (T_l, T_{l+1}]$, $l = 0, \ldots, L-1$ and $d = 1, \ldots, D$, is given by

$$q_{t_{lD+d}}\Big(S(t_{lD+d}) - I(t_{lD+d})\Big).$$

- There is a maximum quantity of gas the buyer can take, which is called the annual contract quantity. Denote the annual contract quantity by ACQ. Similarly, there is a minimum quantity of gas the buyer has to take, which is called the minimum bill. Denote the minimum bill by MB. Both the minimum bill and the annual contract quantity can be violated. If the total gas taken is below the minimum bill or above the annual contract quantity, the buyer has to pay penalties at the end of the contract. More precisely, if $Q_T < MB$ or $Q_T > ACQ$ at time T, there is an out cash flow generated by the penalty, in addition to the cash flow generated by the instant payoff. The possible penalty at time T is given by

$$\mathcal{P}(I, Q_T) = -I \cdot \max\big\{MB - Q_T, 0\big\} - I \cdot \max\big\{Q_T - ACQ, 0\big\}. \quad (4.6)$$

- At time $t_{lD+d} \in (T_l, T_{l+1}]$, $l = 0, \ldots, L-1$ and $d = 1, \ldots, D$, by (4.4) and (4.5), we have the index and the approximated value of the running

average in discrete time as follows:

$$M(t_{lD+d}) = \frac{1}{d}\sum_{j=1}^{d} Z(t_{lD+j}) \text{ for } l = 0, 1, \ldots, L-2. \tag{4.7}$$

$$I(t_{lD+d}) = \begin{cases} K & \text{for } l = 0, \\ \frac{1}{D}\sum_{j=1}^{D} Z(t_{(l-1)D+j}) = M(t_{lD}), & \text{for } l = 1, 2, \ldots, L-1. \end{cases} \tag{4.8}$$

In addition, for t_{lD+d} within the period $(T_l, T_{l+1}]$, that is, $d = 1, 2, \ldots, D-1$, given the values of $M(t_{lD+d})$ and $I(t_{lD+d})$, we have the evaluations of the running average with respect to the oil price and the index, which can be described by

$$M(t_{lD+d+1}) = \frac{d \cdot M(t_{lD+d}) + Z(t_{lD+d+1})}{d+1} \tag{4.9}$$

and

$$I(t_{lD+d+1}) = I(t_{lD+d}), \tag{4.10}$$

respectively.

- The risk-free rate is a fixed constant r.

Remark 4.1. In Chapter 3, the penalty involves a penalty coefficient $\eta \in [0, 1]$ (for instance, see (3.9)). η is the coefficient which controls the functionality of the penalty. In this chapter, we focus on the indexation problem and no longer provide treatment of η. Throughout the rest of this thesis, we simply let $\eta = 1$.

From the buyer's point of view, the goal is to maximize the total expected discounted payoff of the contract, including the penalty. That is, to find the value $V(t_n, S, I, Q)$ where the gas price is S, the index is I and the period to date is Q of this contract at time t_n, which is given by

$$\begin{aligned} V(t_n, S, I, Q) = \sup_{q_{t_k}, n \le k \le N} \mathbb{E}\Bigg[& \sum_{k=n}^{N} e^{-r(t_k - t_n)} q_{t_k} \Big(S(t_k) - I(t_k) \Big) \\ & + e^{-r(T-t_n)} \mathcal{P}\Big(I(t_N), Q_T \Big) \Big| S(t_n) = S, I(t_n) = I, Q_{t_n} = Q \Bigg]. \end{aligned} \tag{4.11}$$

Or, equivalently, for $t_n = t_{lD+d} \in (T_l, T_{l+1}]$, we have

$$V(t_n, S, I, Q) = \sup_{q_{t_m}} \mathbb{E}\Bigg[\underbrace{\sum_{k=d}^{D} e^{-r(t_{lD+k}-t_n)} q_{t_{lD+k}} \Big(S(t_{lD+k}) - I(t_{lD+k}) \Big)}_{\text{part 1}} + \underbrace{\sum_{g=l+1}^{L-1} \sum_{k=1}^{D} e^{-r(t_{gD+k}-t_n)} q_{t_{gD+k}} \Big(S(t_{gD+k}) - I(t_{gD+k}) \Big)}_{\text{part 2}} + \underbrace{e^{-r(T-t_n)} \mathcal{P}\Big(I(t_N), Q_T \Big)}_{\text{part 3}} \Bigg| S(t_n) = S, I(t_n) = I, Q_{t_n} = Q \Bigg],$$

where $m = lD + d, lD + d + 1, \ldots, N$, "part 1" gives the value of all the other instant payoffs in the current month at t_n, "part 2" gives the value of all instant payoffs in all future months at t_n, and "part 3" gives the value of the penalty at t_n.

From (4.11), we have the following result by Bellman's principle of optimality:

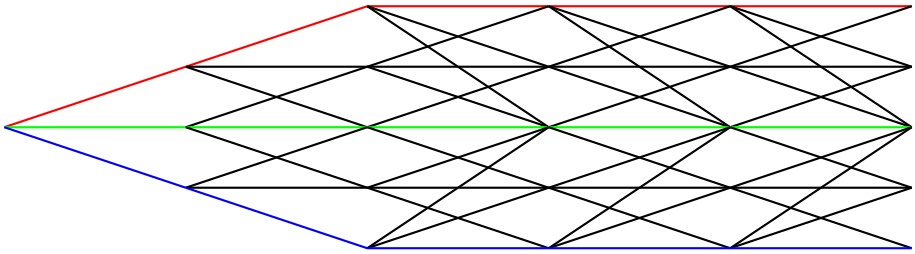
$$\begin{aligned} V(t_n, S, I, Q) = \sup_{q_{t_n}} \Bigg\{ & q_{t_n} \Big(S(t_n) - I(t_n) \Big) \\ & + e^{-r(t_{n+1}-t_n)} \mathbb{E}\Big[V(t_{n+1}, S(t_{n+1}), I(t_{n+1}), Q + q_{t_n}) \Big| \\ & S(t_n) = S, I(t_n) = I, Q_{t_n} = Q \Big] \Bigg\}. \end{aligned} \tag{4.12}$$

In the rest of this chapter, we design an algorithm based on a two-dimensional trinomial tree to compute the expectation of the optimal contract value at time t_{n+1} in (4.12).

4.5 The structure of a trinomial tree

In this chapter, we evaluate the price of the gas sales agreement using a two-dimensional trinomial tree. We first build two separate one-dimensional trinomial trees for the gas price (the gas tree) and the oil price (the oil tree),

FIGURE 4.1: A one-dimensional trinomial tree



respectively. By combining these two one-dimensional trees together, we obtain a two-dimension trinomial tree. The tree building procedures are shown in Section 2.2. Once the tree is built, each node on the 2-D tree can be referenced by a vector of integers (n, s, z), where n indicates that the current time is $n\Delta t$, s and z are the price levels on the gas tree and the oil tree, respectively.

As we can see in Figure 4.1, which shows a typical one-dimensional trinomial tree, there are three possible forms of the tree branching. Denote the highest indexes that can be reached by the gas tree and the oil tree by $s_{\max}$ and $z_{\max}$, respectively. Since both the gas tree and the oil tree are symmetrical, the lowest indexes that can be reached by the gas tree and the oil tree are $-s_{\max}$ and $-z_{\max}$, respectively. Let $S_{n,s}$ be the gas price at time t_n where the price level on the gas tree is s, $s = -s_{\max}, -s_{\max}+1, \ldots, s_{\max}$. Let $Z_{n,z}$ be the oil price at time t_n where the price level on the oil tree is z, $z = -z_{\max}, -z_{\max}+1, \ldots, z_{\max}$. We can see all the forms of the tree branching of both the gas tree and the oil tree in Figure 4.2.

Define the following functions:

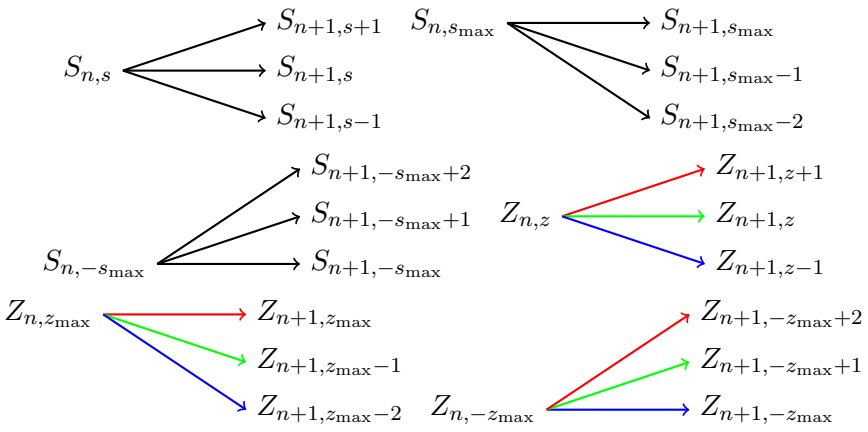
$$\mathfrak{g}(s,b) = \begin{cases} s+b & \text{if} \quad -s_{\max} < s < s_{\max} \\ s+b-1 & \text{if} \quad s = s_{\max} \\ s+b+1 & \text{if} \quad s = -s_{\max} \end{cases}, \tag{4.13}$$

and

$$\mathfrak{h}(z,c) = \begin{cases} z+c & \text{if} \quad -z_{\max} < z < z_{\max} \\ z+c-1 & \text{if} \quad z = z_{\max} \\ z+c+1 & \text{if} \quad z = -z_{\max} \end{cases}, \tag{4.14}$$

where b, $c \in \{-1, 0, 1\}$. Here we use integers -1, 0 and 1 to represent the

FIGURE 4.2: Possible forms of the tree branching for a mean-reverting trinomial tree



lower, middle and upper branches on the trees, respectively. Then $(n+1, \mathfrak{g}(s,b), \mathfrak{h}(z,c))$ gives the node associated with the b branch on the gas tree and the c branch on the oil tree emanating from the node (n,s,z). Now, we define one more function:

$$\mathfrak{j}(n) = \min\{n, z_{\max}\}.$$

As we can see in Figure 4.1, the trinomial tree stops growing at some point. Before the oil tree stops growing, the size of the tree grows over time and n returns the highest level of the oil tree at time t_n. Once the tree stops growing, the size of the tree stops growing and the highest level of the oil tree is simply $z_{\max}$. Hence, $\mathfrak{j}(n)$ gives the highest level of the oil tree at any $t_n \in [0,T]$. Since the oil tree is symmetrical, $-\mathfrak{j}(n)$ gives the lowest level of the oil tree at any $t_n \in [0,T]$. It follows that the total number of nodes on the oil tree at time t_n is $2 \cdot \mathfrak{j}(n) + 1$. Similarly, we can define such a function $\mathfrak{k}(n)$ for the gas tree:

$$\mathfrak{k}(n) = \min\{n, s_{\max}\}.$$

In addition, we denote the probability associated with the b branch on the gas tree and the c branch on the oil tree emanating from node (n,s,z) by $p_{s,z,b,c}$. These probabilities can be computed using (2.27), (2.28) and (2.29).

Remark 4.2. Recall the tree building procedures in Section 2.2. Let ΔX and ΔY be the space steps on the gas fundamental tree and the oil fundamental tree, respectively. That is,

$$\Delta X = \sigma_S \sqrt{3\Delta t} \text{ and } \Delta Y = \sigma_Z \sqrt{3\Delta t}.$$

Let a_n and b_n, $n = 0, 1, \ldots, N$, be the amounts added on each node of the gas fundamental tree and the oil fundamental tree such that the gas tree and the oil tree are consistent with the observed forward curves (see (2.32)), respectively. Then the gas price of level s on the gas tree and the oil price of level z on the oil tree are given by

$$S_{n,s} = e^{s \cdot \Delta X + a_n} \text{ and } Z_{n,z} = e^{z \cdot \Delta Y + b_n}, \tag{4.15}$$

respectively.

4.5.1 The discretization of the running average and the index

To perform the evaluation, at each time t_n, we generate two vectors, $\mathbf{M}_n$ and $\mathbf{I}_n$, in both the running average dimension and the index dimension by discretization, respectively. Since in the first $L-1$ months (i.e. in Month l, $l = 0, 1, \ldots, L-2$), we have the running average variable M, then an integer m which represents the value of the running average should be introduced (we do not need to track the running average in the last month). Similarly, in the last $L-1$ months (i.e. in Month l, $l = 1, 2, \ldots, L-1$), we have the index price which is not a constant, thus we have to add one more integer i which represents the level of the value of the index (we do not need this integer i in the first month where the index is simply a constant K).

4.5.2 The running average vector $\mathbf{M}_n$

$\mathbf{M}_n$ is the vector containing the possible values of the running average M at time t_n. Let $M_{n,m}$ be the value of the running average at time t_n where the level in the running average vector $\mathbf{M}_n$ is m. Let H be the total number of values in $\mathbf{M}_n$, then

$$\mathbf{M}_n = (M_{n,0}, M_{n,1}, \ldots, M_{n,H-1}).$$

We also assume that $\mathbf{M}_n$ is constructed in such a way that

$$M_{n,0} < M_{n,1} < \ldots < M_{n,H-1}.$$

Once the two-dimensional tree is built, we assume that the level m value of the running average at each time t_n, $n = 1, 2, \ldots, N$, is referenced by a deterministic function $\mathfrak{d}_n(m)$ where the value solely depends on m. That is,

$$M_{n,m} = \mathfrak{d}_n(m).$$

Denote the lowest and highest values of the running average M that can be achieved at time t_n through the tree by $M_n^{\min}$ and $M_n^{\max}$, respectively. Referring back to Figure 4.1, $M_n^{\min}$ and $M_n^{\max}$ can be obtained by following the blue branches and red branches, respectively. Then, for $t_{lD+d} \in (T_l, T_{l+1}]$, $d = 1, \ldots, D$ and $l = 0, 1, \ldots, L-2$,

$$M_{lD+d}^{\min} = \frac{1}{d} \sum_{k=1}^{d} Z_{lD+d,-\mathfrak{j}(lD+d)} \tag{4.16}$$

and

$$M_{lD+d}^{\max} = \frac{1}{d} \sum_{k=1}^{d} Z_{lD+d,\mathfrak{j}(lD+d)}. \tag{4.17}$$

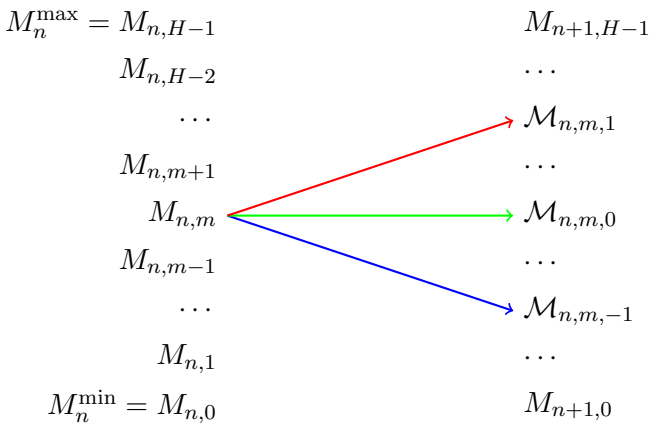
The evaluation of the value of the running average At time t_n, $n = lD + d$, $d = 1, \ldots, D-1$ and $l = 0, 1, \ldots, L-2$, for a given value of the running average $M_{n,m}$, we can get the evaluation of $M_{n,m}$ by following these three possible movements of the oil price Z in Figure 4.2. Note that the evaluation of the running average M only depends on the oil price Z, while the gas price S has no impact on M (see (4.4)). Denote the one time-step forward evaluation of $M_{n,m}$ by $\mathcal{M}_{n,m,c}$, together with (4.9), we have

$$\mathcal{M}_{n,m,c} = \frac{d \cdot M_{n,m} + Z_{n+1,\mathfrak{h}(z,c)}}{d+1}, \tag{4.18}$$

where $c = -1, 0, 1$. Figure 4.3 shows the evaluation of $M_{n,m}$ corresponding to the red, green and blue branches in Figure 4.2.

For the discretization in the running average dimension, we propose two

FIGURE 4.3: One step forward evaluation of the running average



possible approaches: further discretization of the space step on the oil fundamental tree and the non-uniform grid.

Further discretization We use the same discretization method as in Dai, Li and Zhang (2010). Let the space step ΔM in the running average dimension be given by

$$\Delta M = \frac{\Delta Y}{F},$$

where $\Delta Y = \sigma_Z \sqrt{3\Delta t}$ is the space step on the oil fundamental tree. (see Remark 4.2) and F is a positive integer. Then $\mathfrak{d}_n(m)$ is given by

$$\mathfrak{d}_n(m) = M_n^{\min} e^{m\Delta M}. \tag{4.19}$$

Recall that the total number of nodes on the oil tree at time t_n is $2 \cdot \mathrm{j}(n) + 1$, through the further discretization, the total number of values of the running average in the running average vector $\mathbf{M}_n$ is

$$H = 2 \cdot \mathrm{j}(n) \cdot F + 1. \tag{4.20}$$

Non-uniform grid We adopt the non-uniform grid in in't Hout and Foulon (2010). Let M_n, $n = lD + d$, $d = 1, \ldots, D$ and $l = 0, 1, \ldots, L-2$, be given by

$$\mathsf{M}_{lD+d} = \frac{1}{d} \sum_{k=1}^{d} Z_{lD+k,0}. \tag{4.21}$$

That is, M approximates the value of the running average M by assuming that the oil price moves on the oil tree by following the middle branches only (that is, we follow the green branches only in Figure 4.1). Let H be a positive integer and w_n, $n = 1, 2, \ldots, N$, be a series of positive constants. At each time t_n, we first construct a series of equidistant points

$$\xi_0 < \xi_1 < \xi_2 < \ldots < \xi_H,$$

where

$$\xi_m = \sinh^{-1}\Big(\frac{M_n^{\min} - \mathsf{M}_n}{w_n \mathsf{M}_n}\Big) + m\Delta\xi,$$

and

$$\Delta\xi = \frac{1}{H}\Big[\sinh^{-1}\Big(\frac{M_n^{\max} - \mathsf{M}_n}{w_n \mathsf{M}_n}\Big) - \sinh^{-1}\Big(\frac{M_n^{\min} - \mathsf{M}_n}{w_n \mathsf{M}_n}\Big)\Big].$$

Then a non-uniform grid at time t_n in the running average dimension can be obtained through the transformation

$$\begin{aligned} M_{n,m} &= \mathfrak{d}_n(m) \\ &= \mathsf{M}_n + w_n \mathsf{M}_n \sinh(\xi_m). \end{aligned} \tag{4.22}$$

As reported in in't Hout and Foulon (2010), this grid is smooth, and these constants w_n control the fraction of points $M_{n,m}$ that lie in the neighbourhood of M_n. A smaller w_n gives more points in the neighbourhood of M_n. Intuitively speaking, the non-uniform grid could be a better approach than the further discretization if we let more points be in the neighbourhood of M_n. Due to the mean-reverting nature of the oil price, the oil price is unlikely to be too high or too low for a long time period. It follows that value of the running average is even more unlikely to be too large or too small.

4.5.3 The index vector $\mathbf{I}_n$

$\mathbf{I}_n$ is the vector containing the possible values of the index I at time t_n. Let $I_{n,i}$ be the index at time t_n in Month l where the level in the index vector is i. Recall (4.5) and (4.8), the value of the index in Month l, $l = 1, 2, \ldots, L-1$, equals the value of the running average on the last day of Month $l-1$. Then we use the same vectors for $\mathbf{I}_{lD+d}$, $d = 1, \ldots, D$, as the vector for $\mathbf{M}_{lD}$. That is, for $t_{lD+d} \in (T_l, T_{l+1}]$, $d = 1, 2, \ldots, D$ and $l = 1, 2, \ldots, L-1$, we have

$$I_{lD+d,i} = M_{lD,i} = \mathfrak{d}_{lD}(i). \tag{4.23}$$

4.6 Pricing algorithm

In this section, we present an algorithm to price the GSA contract formulated in Section 4.4. This algorithm works backwards in time. Since in the last month there is no further month we do not need to track the running average, which means we also do not need to track the oil price. It follows that the contract value is driven by three variables: the gas price, S, the index, I, and the period to date, Q. In the middle months, i.e. Month 1, Month 2, ..., Month $L-2$, we need to track the running average and the oil price since they will be used to calculate the index in the coming month, which means the contract value is driven by five variables: S, I, Q, the running average, M, and the oil price, Z. In the first month, since the index is a constant K, the contract value is driven by just four variables: S, Q, M and Z.

Based on the above facts, the contract value at time t_n of Month l on the tree can be referenced by

$$\begin{cases} V^0(S_{n,s}, Z_{n,z}, M_{n,m}, Q) & \text{the first month} \\ V^l(S_{n,s}, Z_{n,z}, M_{n,m}, I_{n,i}, Q) & \text{the middle months, i.e. } l = 1, \ldots, L-2 \\ V^{L-1}(S_{n,s}, I_{n,i}, Q) & \text{the last month} \end{cases},$$

where the gas price is $S_{n,s}$, the oil price is $Z_{n,z}$, the value of the running average is $M_{n,m}$, the index is $I_{n,i}$ and the period to date is Q. Similarly, we have

the optimal exercise decisions

$$\begin{cases} q^{0,*}(S_{n,s}, Z_{n,z}, M_{n,m}, Q) & \text{the first month} \\ q^{l,*}(S_{n,s}, Z_{n,z}, M_{n,m}, I_{n,i}, Q) & \text{the middle months, i.e. } l = 1, \ldots, L-2 \\ q^{L-1,*}(S_{n,s}, I_{n,i}, Q) & \text{the last month} \end{cases},$$

corresponding to values of those variables on the tree.

Terminal value Since we have introduced the penalty when the total gas taken during the contract exceeds the annual contract quantity, the terminal optimal decision is different from the case in Section 3.3.1. On the last day of the contract t_N, where $l = L-1$, by letting $q^* = q^{L-1,*}(S_{N,s}, I_{N,i}, Q)$, we have the following scenarios:

- If $S_{N,s} \leq I_{N,i}$,
  - if $Q < MB$, that is the period to date at the terminal is less than the minimum bill. Upon taking a quantity of gas $\breve{q}$ up to that required to avoid the penalty, or the maximum possible, the actual payoff of this decision $\breve{q}$ is

    $$\breve{q} \cdot (S_{N,s} - I_{N,i}) + \breve{q} \cdot I_{N,i} = \breve{q} \cdot S_{N,s} > 0,$$

    where $\breve{q} \cdot I_{N,i}$ is the reduced value of penalty by taking $\breve{q}$. Hence,

    $$q^* = \min\{\max\{MB - Q, q_{\min}\}, q_{\max}\}.$$

  - if $Q \geq MB$, the buyer is under no pressure to meet the minimum bill. Since $S_{N,s} \leq I_{N,i}$, the optimal decision is to take the daily minimum so to reduce the loss from the instant payoff. The daily minimum also avoids or minimizes the penalty from taking more gas than the annual contract quantity. That is,

    $$q^* = q_{\min}.$$

- If $S_{N,s} > I_{N,i}$,

– if $Q + q_{\max} \leq ACQ$, the buyer can safely take the daily maximum to maximize the profit of the instant payoff without worrying about the penalty arising from violating the annual contract quantity. Then,

$$q^* = q_{\max}.$$

– if $0 < ACQ - Q < q_{\max}$, it means the buyer can take a quantity of gas less than the daily maximum in order to gain the profit of the instant payoff without worrying about the penalty. In addition, he/she has the choice of further taking gas and making profit from the instant payoff and increasing the penalty at the same time. Suppose the buyer further takes a quantity of gas $\breve{q}$, then the actual payoff of this decision is

$$\breve{q} \cdot (S_{N,s} - I_{N,i}) - \breve{q} \cdot I_{N,i} = \breve{q} \cdot (S_{N,s} - 2I_{N,i}), \tag{4.24}$$

where $\breve{q} \cdot I_{N,i}$ is the increased value of penalty by further taking $\breve{q}$. Then, we have the optimal decision in this scenario:

$$q^* = \begin{cases} q_{\max} & \text{if } S_{N,s} > 2I_{N,i}, \\ \max\Big\{ \min\big\{ACQ - Q, q_{\max}\big\}, q_{\min}\Big\} & \text{if } S_{N,s} \leq 2I_{N,i}. \end{cases} \tag{4.25}$$

– if $Q \geq ACQ$, this means that, by taking a quantity of gas, the buyer is profiting from the instant payoff and losing money from the penalty at the same time. That is, upon taking $\breve{q}$, the actual payoff of this decision is again given by (4.24), and the optimal decision is also given by (4.25).

In summary, we have the optimal decision at the last day of the contract which is given by:

$$q^* = \begin{cases} \min\Big\{ \max\big\{MB - Q, q_{\min}\big\}, q_{\max}\Big\} & \text{if } S_{N,s} \leq I_{N,i} \\ \max\Big\{ \min\big\{ACQ - Q, q_{\max}\big\}, q_{\min}\Big\} & \text{if } I_{N,i} < S_{N,s} \leq 2I_{N,i} \\ q_{\max} & \text{if } S_{N,s} > 2I_{N,i} \end{cases}. \tag{4.26}$$

Then the contract value on the last day is given by:

$$V^{L-1}(S_{N,s}, Z_{N,z}, I_{N,i}, Q) = q^* \cdot (S_{N,s} - I_{N,i}) + \mathcal{P}\left(I_{N,i}, Q + q^*\right).$$

Remark 4.3. Recall Remark 2.1. In this thesis, we mainly focus on the GSAs introduced in Breslin et al. (2008) where $ACQ = N \cdot q_{\max}$. That is, we mainly focus on the penalty paid when the total volume of gas taken is below the minimum bill. However, for the purpose of completeness, we have also provided the optimal terminal decision for GSAs where $ACQ \leq N \cdot q_{\max}$.

Days within the last month For $n = (L-1)D + d$, $d = 1, \ldots, D-1$, we find the optimal exercise decision by

$$\begin{aligned} q^{L-1,*}(S_{n,s}, I_{n,i}, Q) =& \text{argmax}_{q_{t_n} \in [q_{\min}, q_{\max}]} \Big\{ q_{t_n}(S_{n,s} - I_{n,i}) + \\ & \sum_{b=-1}^{1} p_{s,b} e^{-r\Delta t} V^{L-1}(S_{n+1,\mathfrak{g}(s,b)}, I_{n,i}, Q + q_{t_n}) \Big\}, \end{aligned}$$

where $p_{s,b}$ is the probability associated with the b branch emanating from the node (n, s) on the gas tree. $p_{s,b}$ can be computed by using (2.23), (2.24) and (2.25). By letting $q^* = q^{L-1*}(S_{n,s}, I_{n,i}, Q)$, we have the contract value:

$$\begin{aligned} V^{L-1}(S_{n,s}, I_{n,i}, Q) =& q^*(S_{n,s} - I_{n,i}) + \\ & \sum_{b=-1}^{1} p_{s,b} e^{-r\Delta t} V^{L-1}(S_{n+1,\mathfrak{g}(s,b)}, I_{n,i}, Q + q^*). \end{aligned} \tag{4.27}$$

Days within middle months Similarly, for $n = lD + d$, $d = 1, \ldots, D-1$ and $l = 1, 2, \ldots, L-2$, we find the optimal exercise decision by

$$\begin{aligned} q^{l,*}(S_{n,s}, Z_{n,z}, M_{n,m}, I_{n,i}, Q) =& \text{argmax}_{q_{t_n} \in [q_{\min}, q_{\max}]} \Big\{ q_{t_n}(S_{n,s} - I_{n,i}) + \\ & \sum_{b,c=-1}^{1} p_{s,z,b,c} e^{-r\Delta t} \cdot \\ & V^{l}(S_{n+1,\mathfrak{g}(s,b)}, Z_{n+1,\mathfrak{h}(z,c)}, \mathcal{M}_{n,m,c}, I_{n,i}, Q + q_{t_n}) \Big\}. \end{aligned} \tag{4.28}$$

By letting $q^* = q^{l,*}(S_{n,s}, Z_{n,z}, M_{n,m}, I_{n,i}, Q)$, we have the contract value

$$\begin{aligned} V^l(S_{n,s}, Z_{n,z}, M_{n,m}, I_{n,i}, Q) =& q^*(S_{n,s} - I_{n,i}) + \\ & \sum_{b,c=-1}^{1} p_{s,z,b,c} e^{-r\Delta t} \cdot \\ & V^l(S_{n+1,\mathfrak{g}(s,b)}, Z_{n+1,\mathfrak{h}(z,c)}, \mathcal{M}_{n,m,c}, I_{n,i}, Q + q^*). \end{aligned} \tag{4.29}$$

Days within the first month Again, for $n = d$, $d = 1, \ldots, D-1$, we find the optimal exercise decision by

$$\begin{aligned} q^{0,*}(S_{n,s}, Z_{n,z}, M_{n,m}, Q) =& \text{argmax}_{q_{t_n} \in [q_{\min}, q_{\max}]} \Bigg\{ q_{t_n}(S_{n,s} - K) + \\ & \sum_{b,c=-1}^{1} p_{s,z,b,c} e^{-r\Delta t} \cdot \\ & V^0(S_{n+1,\mathfrak{g}(s,b)}, Z_{n+1,\mathfrak{h}(z,c)}, \mathcal{M}_{n,m,c}, Q + q_{t_n}) \Bigg\}. \end{aligned} \tag{4.30}$$

By letting $q^* = q^{0,*}(S_{n,s}, Z_{n,z}, M_{n,m}, Q)$, we have the contract value

$$\begin{aligned} V^0(S_{n,s}, Z_{n,z}, M_{n,m}, Q) =& q^*(S_{n,s} - K) + \\ & \sum_{b,c=-1}^{1} p_{s,z,b,c} e^{-r\Delta t} \cdot \\ & V^0(S_{n+1,\mathfrak{g}(s,b)}, Z_{n+1,\mathfrak{h}(z,c)}, \mathcal{M}_{n,m,c}, Q + q^*). \end{aligned} \tag{4.31}$$

Linear interpolation Recall in Section 4.5, we have a discretized grid in the running average dimension. When

$$V^l(S_{n+1,\mathfrak{g}(s,b)}, Z_{n+1,\mathfrak{h}(z,c)}, \mathcal{M}_{n,m,c}, I_{n,i}, Q + q_{t_n})$$

or

$$V^0(S_{n+1,\mathfrak{g}(s,b)}, Z_{n+1,\mathfrak{h}(z,c)}, \mathcal{M}_{n,m,c}, Q + q_{t_n})$$

in (4.28) and (4.30) (for simplicity, we use $V(\mathcal{M}_{n,m,b})$) does not fall exactly on the discrete nodes of the grid, we use linear interpolation:

$$V(\mathcal{M}_{n,m,b}) = V(M_{n+1,\bar{m}}) + (\mathcal{M}_{n,m,b} - M_{n+1,\bar{m}})\frac{V(M_{n+1,\bar{m}+1}) - V(M_{n+1,\bar{m}})}{M_{n+1,\bar{m}+1} - M_{n+1,\bar{m}}}, \tag{4.32}$$

where $\bar{m}$ is an integer such that $M_{n+1,\bar{m}}$ is the largest value in the running average vector $\mathbf{M}_{n+1}$ less than or equal to $\mathcal{M}_{n,m,b}$.

Matching point At time t_{lD}, $l = 2, \ldots, L-2$, we are on the last day of Month $l-1$. Given the current value of the running average $M(t_{lD})$, the index for time t_{lD+1}, which is the first day of Month l, is $I(t_{lD+1}) = M(t_{lD})$ (see (4.8)). In addition, at time t_{lD+1}, $l = 2, \ldots, L-2$, by (4.7), the value of the running average is

$$M(t_{lD+1}) = \frac{1}{1}\sum_{j=1}^{1} Z(t_{lD+j}) = Z(t_{lD+1}). \tag{4.33}$$

That is, at time t_{lD}, if the contract value $V^l(S_{lD,s}, Z_{lD,z}, M_{lD,m}, I_{n,i}, Q)$ is given, upon taking the volume $q_{t_{lD}}$, the contract value corresponding to the b branch on the gas tree and c branch on the oil tree is

$$V^l(S_{lD+1,\mathfrak{g}(s,b)}, Z_{lD+1,\mathfrak{h}(z,c)}, Z_{lD+1,\mathfrak{h}(z,c)}, M_{lD,m}, Q + q_{t_{lD}}).$$

Then we can find the optimal decision $q^{l*}(S_{lD,s}, Z_{lD,z}, M_{lD,m}, I_{lD,i}, Q)$ by

$$\begin{aligned}
&q^{l,*}(S_{lD,s}, Z_{lD,z}, M_{lD,m}, I_{lD,i}, Q)\\
=&\text{argmax}_{q_{t_{lD}} \in [q_{\min}, q_{\max}]} \Bigg\{ q_{t_{lD}}(S_{lD,s} - I_{lD,i}) +\\
&\sum_{b,c=-1}^{1} p_{s,z,b,c} e^{-r\Delta t} V^{l+1}(S_{lD+1,\mathfrak{g}(s,b)}, Z_{lD+1,\mathfrak{h}(z,c)}, Z_{lD+1,\mathfrak{h}(z,c)}, M_{lD,m}, Q + q_{t_{lD}}) \Bigg\}.
\end{aligned} \tag{4.34}$$

By letting $q^* = q^{l,*}(S_{lD,s}, Z_{lD,z}, M_{lD,m}, I_{lD,i}, Q)$, we have the contract value

$$\begin{aligned}
&V^l(S_{lD,s}, Z_{lD,z}, M_{lD,m}, I_{lD,i}, Q)\\
=&q^*(S_{lD,s} - I_{lD,i}) +\\
&\sum_{b,c=-1}^{1} p_{s,z,b,c} e^{-r\Delta t} V^{l+1}(S_{lD+1,\mathfrak{g}(s,b)}, Z_{lD+1,\mathfrak{h}(z,c)}, Z_{lD+1,\mathfrak{h}(z,c)}, M_{lD,m}, Q + q^*).
\end{aligned} \tag{4.35}$$

Similarly, we have the optimal decision and contract value on the last day of Month 0, that is, at time t_D:

$$\begin{aligned}
&q^{0,*}(S_{D,s}, Z_{D,z}, M_{D,m}, Q)\\
=&\text{argmax}_{q_{t_D} \in [q_{\min}, q_{\max}]} \Bigg\{ q_{t_D}(S_{D,s} - K) +\\
&\sum_{b,c=-1}^{1} p_{s,z,b,c} e^{-r\Delta t} V^1(S_{D+1,\mathfrak{g}(s,b)}, Z_{D+1,\mathfrak{h}(z,c)}, Z_{D+1,\mathfrak{h}(z,c)}, M_{D,m}, Q + q_{t_D}) \Bigg\},
\end{aligned}$$

and with $q^* = q^{0,*}(S_{D,s}, Z_{D,z}, M_{D,m}, Q)$, we have

$$\begin{aligned}
&V^0(S_{D,s}, Z_{D,z}, M_{D,m}, Q)\\
=&q^*(S_{D,s} - K) +\\
&\sum_{b,c=-1}^{1} p_{s,z,b,c} e^{-r\Delta t} V^1(S_{D+1,\mathfrak{g}(s,b)}, Z_{D+1,\mathfrak{h}(z,c)}, Z_{D+1,\mathfrak{h}(z,c)}, M_{D,m}, Q + q^*).
\end{aligned}$$

Also, we have the optimal decision and contract value on the last day of Month $L-2$, that is, at time $t_{(L-1)D}$:

$$\begin{aligned}
&q^{L-2,*}(S_{(L-1)D,s}, Z_{(L-1)D,z}, M_{(L-1)D,m}, I_{(L-1)D,i}, Q)\\
=&\text{argmax}_{q_{t_{(L-1)D}} \in [q_{\min}, q_{\max}]} \Bigg\{ q_{t_{(L-1)D}}(S_{(L-1)D,s} - I_{(L-1)D,i}) +\\
&\sum_{b,c=-1}^{1} p_{s,z,b,c} e^{-r\Delta t} V^{L-1}(S_{(L-1)D+1,\mathfrak{g}(s,b)}, M_{(L-1)D,m}, Q + q_{t_{(L-1)D}}) \Bigg\},
\end{aligned}$$

and with $q^* = q^{L-2,*}(S_{(L-1)D,s}, Z_{(L-1)D,z}, M_{(L-1)D,m}, I_{(L-1)D,i}, Q)$,

$$\begin{aligned}
&V^{L-2}(S_{(L-1)D,s}, Z_{(L-1)D,z}, M_{(L-1)D,m}, I_{(L-1)D,i}, Q) \\
=&q^*(S_{(L-1)D,s} - I_{(L-1)D,i}) + \\
&\sum_{b,c=-1}^{1} p_{s,z,b,c} e^{-r\Delta t} V^{L-1}(S_{(L-1)D+1,\mathfrak{g}(s,b)}, M_{(L-1)D,m}, Q + q^*).
\end{aligned}$$

Remark 4.4. Based on our framework, for a function which is a bit smoother than (4.6), Theorem 2.2 still applies. In our case, however, (4.6) is clearly not differentiable at some points with respect to Q. However, the bang-bang consumption is observed in our numerical examples. Based on our numerical observation, we make the following assumption: the GSAs in this chapter have the so-called bang-bang consumption, i.e. the optimal decision is either the daily maximum $q_{\max}$ or the daily minimum $q_{\min}$. Without loss of generality, we further assume that $q_{\max} = \bar{q}$ and $q_{\min} = 0$ in the rest of this chapter.

Remark 4.5. In (4.29), we need to find the contract values of all possible combinations of $(S_{n,s}, Z_{n,z}, M_{n,m}, I_{n,i}, Q)$. Recall that, at time t_n, the numbers of nodes on the gas tree and the oil tree are $2 \cdot \mathfrak{k}(n) + 1$ and $2 \cdot \mathfrak{j}(n) + 1$, respectively. At the same time, the numbers of values in the running average vector $\mathbf{M}_n$ and the index vector $\mathbf{I}_n$ are both H. Under the bang-bang consumption, according to Remark 2.6, if we let $q_{\max} = 1$ and $q_{\min} = 0$, the possible periods to date at time t_n is $0, 1, \ldots, n-1$, then it is not hard to see that (4.29) has complexity of the order of

$$\mathcal{O}\Big(n \cdot H^2 \cdot \Big(2 \cdot \mathfrak{k}(n) + 1\Big)\Big(2 \cdot \mathfrak{j}(n) + 1\Big)\Big).$$

Similarly, (4.31) has complexity of the order of

$$\mathcal{O}\Big(n \cdot H \cdot \Big(2 \cdot \mathfrak{k}(n) + 1\Big)\Big(2 \cdot \mathfrak{j}(n) + 1\Big)\Big),$$

and (4.27) has complexity of the order of

$$\mathcal{O}\Big(n \cdot H \cdot \Big(2 \cdot \mathfrak{k}(n) + 1\Big)\Big).$$

$\alpha_S = 0.8$	$\sigma_S = 0.5$	$\alpha_I = 0.6$	$\sigma_I = 0.6$	$r = 0.05$
$\rho = 0.5$	$q_{\min} = 0$	$q_{\max} = 1$	$MB = 10$	$ACQ = 16$
$T = 1$	$N = 16$	$D = 4$	$L = 4$	

TABLE 4.1: Parameter values.

4.7 A simple example

In this section, to better illustrate our algorithm, we consider a simple GSA.

The implementation starts by building a two-dimensional trinomial tree. To keep this example simple, we use the parameters in Table 4.1 so that the size of the generated tree is small. In addition, we let the observed forward curves for gas and oil be

$$F^S(0,t_n) = F^Z(0,t_n) = 100 + 50 \cdot \sin\Big(2\pi\frac{n}{N}\Big).$$

Table 4.2 shows the gas prices on the gas tree and Table 4.3 shows the oil prices on the oil tree.

When generating the running average vector $\mathbf{M}_n$ at time t_n, we use the non-uniform grid approach with the following parameters:

$$H = 15, \quad w_n = \frac{1}{3} \text{ for } n = 1, \ldots, N.$$

To obtain such a grid, we also need to find $M_n^{\max}$, $M_n^{\min}$ and M_n. For instance, on Day 7, by Figure 4.1 and (4.21),

$$\begin{aligned}
M_7^{\max} =& \frac{Z_{5,5} + Z_{6,5} + Z_{7,5}}{3} = \frac{510.513 + 469.049 + 409.913}{3} = 463.158 \\
=& M_{7,14}, \\
M_7^{\min} =& \frac{Z_{5,-5} + Z_{6,-5} + Z_{7,-5}}{3} = \frac{37.9906 + 34.9051 + 30.5043}{3} = 34.4667 \\
=& M_{7,0}, \\
\mathsf{M}_7 =& \frac{Z_{5,0} + Z_{6,0} + Z_{7,0}}{3} = \frac{139.265 + 127.954 + 111.822}{3} = 136.347.
\end{aligned}$$

TABLE 4.2: The gas tree.

Levels	Day 0	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13	Day 14	Day 15	Day 16
Level 4					347.127	336.571	310.164	271.844	227.317	183.192	145.992	121.176	112.321	120.597	144.593	180.542	222.89
Level 3				274.027	279.551	271.05	249.784	218.924	183.064	147.529	117.571	97.586	90.4551	97.1198	116.445	145.395	179.499
Level 2			205.625	220.681	225.13	218.284	201.158	176.305	147.427	118.809	94.6835	78.5887	72.8459	78.2132	93.7761	117.091	144.556
Level 1		146.781	165.595	177.721	181.303	175.79	161.998	141.984	118.727	95.6805	76.2513	63.2896	58.6649	62.9872	75.5205	94.2965	116.415
Level 0	100	118.207	133.358	143.124	146.009	141.568	130.461	114.343	95.614	77.0541	61.4072	50.9689	47.2444	50.7253	60.8187	75.9396	93.752
Level -1		95.1954	107.397	115.261	117.585	114.009	105.064	92.0837	77.0005	62.0538	49.4529	41.0466	38.0472	40.8505	48.979	61.1562	75.501
Level -2			86.4899	92.8231	94.6943	91.8145	84.6109	74.1576	62.0107	49.9736	39.8258	33.056	30.6405	32.898	39.4441	49.2508	60.8031
Level -3				74.753	76.2599	73.9407	68.1395	59.7211	49.9389	40.2451	32.0728	26.6209	24.6756	26.4937	31.7654	39.663	48.9664
Level -4					61.4142	59.5465	54.8746	48.095	40.2171	32.4105	25.8291	21.4385	19.8719	21.3361	25.5816	31.9417	39.434

TABLE 4.3: The oil tree

Levels	Day 0	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13	Day 14	Day 15	Day 16
Level 5						510.513	469.049	409.913	341.818	274.733	218.386	180.822	167.219	179.142	214.337	267.091	329.116
Level 4					407.318	393.707	361.731	316.125	263.61	211.874	168.419	139.45	128.959	138.155	165.296	205.981	253.815
Level 3				308.904	314.124	303.627	278.967	243.795	203.296	163.397	129.885	107.544	99.4534	106.545	127.477	158.852	195.742
Level 2			222.705	238.227	242.252	234.157	215.139	188.015	156.782	126.012	100.167	82.9377	76.6985	82.1674	98.31	122.507	150.956
Level 1		152.751	171.75	183.72	186.825	180.582	165.915	144.997	120.91	97.1805	77.249	63.9616	59.1499	63.3675	75.8166	94.4773	116.417
Level 0	100	117.801	132.453	141.685	144.079	139.265	127.954	111.822	93.2459	74.9456	59.5744	49.3271	45.6164	48.869	58.4698	72.8609	89.781
Level -1		90.8484	102.148	109.268	111.114	107.401	98.6781	86.237	71.9113	57.798	45.9438	38.0411	35.1793	37.6878	45.0919	56.1903	69.2391
Level -2			78.7766	84.2671	85.6911	82.8276	76.1005	66.506	55.458	44.5738	35.4319	29.3373	27.1303	29.0648	34.7749	43.334	53.3972
Level -3				64.9868	66.085	63.8767	58.6887	51.2894	42.7692	34.3753	27.325	22.6249	20.9229	22.4148	26.8184	33.4192	41.1799
Level -4					50.9647	49.2617	45.2607	39.5544	32.9836	26.5102	21.0731	17.4483	16.1357	17.2863	20.6823	25.7729	31.7579
Level -5						37.9906	34.9051	30.5043	25.4369	20.4447	16.2515	13.4562	12.4439	13.3312	15.9502	19.876	24.4917

TABLE 4.4: The running average vectors.

Levels	Day 0	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13	Day 14	Day 15	Day 16
Level 0	N/A	90.8484	84.8125	78.2039	71.3941	37.9906	36.4478	34.4667	32.2092	20.4447	18.3481	16.7175	15.6491	N/A	N/A	N/A	N/A
Level 1	N/A	117.801	92.9145	89.8764	86.5195	49.2617	65.1648	61.6227	57.5866	26.5102	32.8045	29.889	27.9789	N/A	N/A	N/A	N/A
Level 2	N/A	152.751	100.323	100.178	99.4683	63.8767	87.3796	82.63	77.2181	34.3753	43.9876	40.0783	37.5169	N/A	N/A	N/A	N/A
Level 3	N/A	N/A	107.197	109.454	110.834	82.8276	105.203	99.4842	92.9684	44.5738	52.9599	48.2531	45.1693	N/A	N/A	N/A	N/A
Level 4	N/A	N/A	113.684	118.017	121.138	107.401	120.327	113.787	106.334	57.798	60.5736	55.1902	51.6631	N/A	N/A	N/A	N/A
Level 5	N/A	N/A	119.925	126.156	130.852	139.265	134.19	126.896	118.585	74.9456	67.5522	61.5486	57.6151	N/A	N/A	N/A	N/A
Level 6	N/A	N/A	126.055	134.143	140.421	180.582	148.108	140.057	130.884	97.1805	74.5585	67.9322	63.5907	N/A	N/A	N/A	N/A
Level 7	N/A	N/A	132.204	142.248	150.285	234.157	163.403	154.521	144.4	126.012	82.2582	74.9476	70.1578	N/A	N/A	N/A	N/A
Level 8	N/A	N/A	138.505	150.744	160.895	303.627	181.528	171.661	160.418	163.397	91.3826	83.2612	77.94	N/A	N/A	N/A	N/A
Level 9	N/A	N/A	145.095	159.915	172.737	393.707	204.206	193.106	180.458	211.874	102.799	93.6627	87.6768	N/A	N/A	N/A	N/A
Level 10	N/A	N/A	152.115	170.071	186.354	510.513	233.59	220.893	206.425	274.733	117.591	107.14	100.293	N/A	N/A	N/A	N/A
Level 11	N/A	N/A	159.716	181.553	202.371	N/A	272.472	257.662	240.786	N/A	137.165	124.974	116.987	N/A	N/A	N/A	N/A
Level 12	N/A	N/A	168.062	194.747	221.522	N/A	324.546	306.905	286.804	N/A	163.379	148.859	139.346	N/A	N/A	N/A	N/A
Level 13	N/A	N/A	177.333	210.097	244.684	N/A	394.759	373.302	348.852	N/A	198.725	181.064	169.492	N/A	N/A	N/A	N/A
Level 14	N/A	N/A	187.728	228.12	272.919	N/A	489.781	463.158	432.823	N/A	246.56	224.647	210.29	N/A	N/A	N/A	N/A

TABLE 4.5: The index vectors.

Levels	Day 0	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13	Day 14	Day 15	Day 16
Level 0	N/A	N/A	N/A	N/A	N/A	71.3941	71.3941	71.3941	71.3941	32.2092	32.2092	32.2092	32.2092	15.6491	15.6491	15.6491	15.6491
Level 1	N/A	N/A	N/A	N/A	N/A	86.5195	86.5195	86.5195	86.5195	57.5866	57.5866	57.5866	57.5866	27.9789	27.9789	27.9789	27.9789
Level 2	N/A	N/A	N/A	N/A	N/A	99.4683	99.4683	99.4683	99.4683	77.2181	77.2181	77.2181	77.2181	37.5169	37.5169	37.5169	37.5169
Level 3	N/A	N/A	N/A	N/A	N/A	110.834	110.834	110.834	110.834	92.9684	92.9684	92.9684	92.9684	45.1693	45.1693	45.1693	45.1693
Level 4	N/A	N/A	N/A	N/A	N/A	121.138	121.138	121.138	121.138	106.334	106.334	106.334	106.334	51.6631	51.6631	51.6631	51.6631
Level 5	N/A	N/A	N/A	N/A	N/A	130.852	130.852	130.852	130.852	118.585	118.585	118.585	118.585	57.6151	57.6151	57.6151	57.6151
Level 6	N/A	N/A	N/A	N/A	N/A	140.421	140.421	140.421	140.421	130.884	130.884	130.884	130.884	63.5907	63.5907	63.5907	63.5907
Level 7	N/A	N/A	N/A	N/A	N/A	150.285	150.285	150.285	150.285	144.4	144.4	144.4	144.4	70.1578	70.1578	70.1578	70.1578
Level 8	N/A	N/A	N/A	N/A	N/A	160.895	160.895	160.895	160.895	160.418	160.418	160.418	160.418	77.94	77.94	77.94	77.94
Level 9	N/A	N/A	N/A	N/A	N/A	172.737	172.737	172.737	172.737	180.458	180.458	180.458	180.458	87.6768	87.6768	87.6768	87.6768
Level 10	N/A	N/A	N/A	N/A	N/A	186.354	186.354	186.354	186.354	206.425	206.425	206.425	206.425	100.293	100.293	100.293	100.293
Level 11	N/A	N/A	N/A	N/A	N/A	202.371	202.371	202.371	202.371	240.786	240.786	240.786	240.786	116.987	116.987	116.987	116.987
Level 12	N/A	N/A	N/A	N/A	N/A	221.522	221.522	221.522	221.522	286.804	286.804	286.804	286.804	139.346	139.346	139.346	139.346
Level 13	N/A	N/A	N/A	N/A	N/A	244.684	244.684	244.684	244.684	348.852	348.852	348.852	348.852	169.492	169.492	169.492	169.492
Level 14	N/A	N/A	N/A	N/A	N/A	272.919	272.919	272.919	272.919	432.823	432.823	432.823	432.823	210.29	210.29	210.29	210.29

Table 4.4 shows the resulting running average vectors $\mathbf{M}_n$ for $n = 1, 2, \ldots, 12$. Days 1, 5 and 9 are the first days in Month 0, Month 1 and Month 2, respectively. As we can see in Table 4.4, according to (4.33), the values of the running average on Days 1, 5 and 9 are exactly the same as the oil prices (see Table 4.3) on these corresponding days. That is, on the first day of each month, instead of using the non-uniform grid, we use the exact values of the running average, which are simply the oil prices on the same day.

Once we have the running average vectors, (4.23) is used to generate the index vectors in Table 4.5. For instance, the index values on Day 13, 14, 15 and 16 in Table 4.5 are the same as the values of the running average on Day 12 in Table 4.4. Furthermore, we do not need to generate either the running average vectors in Month 3 or the index vectors in Month 0, since they are irrelevant to the contract value (see(4.27) and (4.31)).

4.8 The modelling of GSAs in continuous time

In Edoli (2013) (also see Basei, Cesaroni and Vargiolu (2014) and Edoli, Fiorenzani and Vargiolu (2016)), the author models the GSA contract in continuous time and shows that its value is the solution of a Hamilton–Jacobi–Bellman (HJB) equation. In this section, we present the work in Edoli (2013) so that we can show in Section 4.9 that the algorithm built in Section 4.6 is consistent with the HJB equation.

Our algorithm is based on a two-dimensional trinomial tree. Recall the tree building procedures in Section 2.2. To get such a trinomial tree, we first build a fundamental tree for a two-dimensional Markov process:

$$dX(t) = -\alpha_S X(t)dt + \sigma_S dB^S(t), \tag{4.36}$$

$$dY(t) = -\alpha_Z Y(t)dt + \sigma_Z dB^Z(t), \tag{4.37}$$

which is obtained by assuming $\theta^S(t) = \theta^Z(t) = 0$ in (4.1) and (4.2). Then we shift the nodes on this fundamental tree by adding a proper drift in order to be consistent with the observed forward curve. Recall Remark 4.2, the gas price and the oil price on the trinomial tree are given by

$$S_{n,s} = e^{s\cdot\Delta X + a_n}, \quad Z_{n,z} = e^{z\cdot\Delta Y + b_n}.$$

a_n and b_n (see (2.32)) are the drifts we have added on the fundamental trees to capture the information from the observed forward curves of the gas and oil, respectively. Hence, we make the following assumption:

Assumption 4.1. $a(t)$ and $b(t)$ are continuous real-valued deterministic functions which are continuously differentiable on $[0, T]$ such that

$$S(t) = e^{X(t)+a(t)} \text{ and } Z(t) = e^{Y(t)+b(t)}$$

for $t \in [0, T]$, where S and Z are given by (4.1) and (4.2), respectively. X and Y are given by (4.36) and (4.37), respectively. Furthermore, we assume that

$$a_n = a(t_n) \text{ and } b_n = b(t_n).$$

Then, rigorously speaking, our algorithm is built on the following model:

$$\begin{cases} dX(t) = -\alpha_S X(t)dt + \sigma_S dB^S(t) \\ S(t) = e^{X(t)+a(t)} \\ dY(t) = -\alpha_Z Y(t)dt + \sigma_Z dB^Z(t) \\ Z(t) = e^{Y(t)+b(t)} \end{cases}, \tag{4.38}$$

where $X(0) = Y(0) = 0$. Indeed, this model is a particular case of the model in Schwartz and Smith (2000). By Itô's formula, one can easily derive

$$dS(t) = \Big[\Theta^S(t) - \alpha_S \ln S(t)\Big] S(t)dt + \sigma_S S(t) dB^S(t), \tag{4.39}$$

$$dZ(t) = \Big[\Theta^Z(t) - \alpha_Z \ln Z(t)\Big] Z(t)dt + \sigma_Z Z(t) dB^S(t). \tag{4.40}$$

where

$$\Theta^S(t) = \frac{\partial a(t)}{\partial t} + \alpha_S a(t) + \frac{1}{2}\sigma_S^2,$$

$$\Theta^Z(t) = \frac{\partial b(t)}{\partial t} + \alpha_Z b(t) + \frac{1}{2}\sigma_Z^2.$$

Now, let us recall some notations in Section 4.3 and 4.4 and define some new notations:

- The time horizon $[0, T]$ is equally spaced into L pieces of length $\delta = \frac{T}{L}$.

- $T_l = l \cdot \delta$ is the beginning of the l-th period, $l = 0, 1, \ldots, L-1$ and $T_L = T$.
- $S(t)$, $Z(t)$, $M(t)$ and $I(t)$ are the gas price, oil price, value of the running average and index at time $t \in [0, T]$, respectively. In the rest of this chapter, $S(t)$ and $Z(t)$ are given by (4.39) and (4.40), respectively. $M(t)$ and $I(t)$ are still given by (4.4) and (4.3), respectively.
- Denote $q(t)$ the amount of gas taken in every sub-period, and $0 \leq q(t) \leq \bar{q}$.
- Let $Q(t)$ be the cumulative amount of gas taken up to time t (the period to date),
$$Q(t) = \int_0^t q(u)du. \tag{4.41}$$
It follows that $dQ(t) = q(t)dt$.
- MB is the minimum bill and ACQ is the annual contract quantity.

In the same sprit of Section 4.4, we have the contract value of a GSA at time 0, which is given by

$$V(0) = \sup_{q(u)} \mathbb{E}\left[\int_0^T e^{-ru} q(u)\big(S(u) - I(u)\big)du + e^{-rT}\mathcal{P}\Big(I(T), Q(T)\Big)\right] \tag{4.42}$$

where $\mathcal{P}\Big(I(T), Q(T)\Big)$ is given by (4.6).

Remark 4.6. By the nature of the index (see equation (4.5)), the value of I does not change continuously, it only changes with a jump at each time T_l, $t = 1, 2, \ldots, L-1$. In addition, since the index at time $t \in (0, T_1]$ is a constant, K, and the index I at time $t \in (T_l, T_{l+1}]$, $l = 1, \ldots, L-1$, is calculated by using the oil prices which are already realized, we assume that in each period $(T_l, T_{l+1}]$, $l = 0, \ldots, L-1$, the index I is known and fixed. That is, in each period $(T_l, T_{l+1}]$, we have $dI(t) = 0$.

Now, we introduce the following notations:

- $V^{\text{tml}}(T, S, I, Q)$ is the penalty at terminal T where the gas price is S, the index is I and the period to date is Q.
- $V^{L-1}(t, S, I, Q)$ is the contract value for $t \in [T_{L-1}, T_L]$ where the gas price is S, the index is I and the period to date is Q.

- $V^l(t,S,Z,M,I,Q)$ is the contract value for $t \in (T_l, T_{l+1}]$, $l = 0,1,\ldots,L-2$, where the gas price is S, the oil price is Z, the index is I and the period to date is Q.

Then, the value function (4.42) can also be defined on months by the Dynamic Programming Principle, which gives

$$V^{\text{tml}}(T,S,I,Q) = \mathcal{P}\big(I(T),Q(T)\big), \tag{4.43}$$

$$\begin{aligned} V^{L-1}(t,S,I,Q) = \sup_{q(u)} \mathbb{E}\Bigg[& \int_t^{T_L} e^{-r(u-t)} q(u)\big(S(u)-I\big)du \\ & + e^{-r(T_L-t)} V^{\text{tml}}(T,S,I,Q)\Big| S(t)=S, Q(t)=Q\Bigg] \end{aligned}$$

$$\text{for } t \in [T_{L-1}, T_L], \tag{4.44}$$

$$\begin{aligned} V^l(t,S,Z,M,I,Q) = \sup_{q(u)} \mathbb{E}\Bigg[& \int_t^{T_{l+1}} e^{-r(u-t)} q(u)\big(S(u)-I\big)du \\ & + e^{-r(T_{l+1}-t)} \Psi^l(S,Z,M,Q)\Big| S(t)=S, Z(t)=Z, \\ & M(t)=M, Q(t)=Q\Bigg] \quad \text{for } t \in (T_l, T_{l+1}], \end{aligned}$$

$$l = 0,1,\ldots,L-2. \tag{4.45}$$

In equations (4.44) and (4.45), the part $q(u)\big(S(u)-I\big)$ gives the instant payoff upon taking the volume of gas $q(u)$ at time u. By multiplying the discount factor $e^{-r(u-t)}$, it gives the value of this instant payoff at time t. The integral from t to T_{l+1} gives the value of the sum of all instant payoffs from t to T_{l+1} at time t. In equation (4.43), $\mathcal{P}\big(I,Q(T)\big)$ is the penalty function. Note that, due to the daily constraints, the value of the period to date at time T is limited on the interval $[0,T]$. That is,

$$0 \le Q(T) \le \bar{q}T.$$

In terms of the part $\Psi^l(x)$, $l = 0,1,\ldots,L-2$, we have the following two scenarios:

- When $l = 0,1,\ldots,L-3$, we find $\Psi^l(x)$ at time T_{l+1} by using V^{l+1}. Let $T_l^+ = T_l + \varepsilon$, $l = 0,1,\ldots,L-3$. ε is a constant and $\varepsilon \ll \delta$. By the fact

that

$$\begin{aligned}\lim_{\varepsilon\to 0} M(T_{l+1}^{+}) &= \lim_{\varepsilon\to 0} \frac{1}{T_{l+1}^{+} - T_{l+1}} \int_{T_{l+1}}^{T_{l+1}^{+}} Z(u)du \\ &= \lim_{\varepsilon\to 0} \frac{1}{T_{l+1} + \varepsilon - T_{l+1}} \int_{T_{l+1}}^{T_{l+1}+\varepsilon} Z(u)du \\ &= Z(T_{l+1}),\end{aligned}$$

and

$$\lim_{\varepsilon\to 0} S(T_{l+1}^{+}) = S(T_{l+1}), \qquad \lim_{\varepsilon\to 0} Z(T_{l+1}^{+}) = Z(T_{l+1}),$$

together with (4.5), we have

$$\begin{aligned}\Psi^{l}(S, Z, M, Q) &= V^{l+1}(T_{l+1}^{+}, S, Z, M(T_{l+1}^{+}), M, Q) \\ &= V^{l+1}(T_{l+1}^{+}, S, Z, Z, M, Q),\end{aligned} \tag{4.46}$$

where, from now on, $T_{l+1}^{+} = \lim_{\varepsilon\to 0}(T_{l+1} + \varepsilon)$.

- When $l = L - 2$, by (4.5), $\Psi^{l}(x)$ is simply

$$\Psi^{L-2}(S, Z, M, Q) = V^{L-1}(T_{L-1}, S, M, Q). \tag{4.47}$$

The combination of (4.46) and (4.47) gives

$$\Psi^{l}(S, Z, M, Q) = \begin{cases} V^{L-1}(T_{L-1}, S, M, Q) & \text{for } l = L-2, \\ V^{l+1}(T_{l+1}^{+}, S, Z, Z, M, Q), & \text{for } l = 0, 1, \ldots, L-3. \end{cases} \tag{4.48}$$

It has been proved in Edoli (2013) that, for this type of problem, the value function (4.44) satisfies

$$\begin{aligned}\frac{\partial V^{L-1}}{\partial t} - rV^{L-1} + [\Theta^{S}(t) - \alpha_S \ln S]S\frac{\partial V^{L-1}}{\partial S} + \frac{1}{2}\sigma_S^2 S^2 \frac{\partial^2 V^{L-1}}{\partial S^2} \\ + \max_{q\in[0,\bar{q}]}\left\{ q\Big[(S - I) + \frac{\partial V^{L-1}}{\partial Q}\Big]\right\} = 0\end{aligned} \tag{4.49}$$

with the terminal condition

$$V^{L-1}(T_L, S, I, Q) = V^{\text{tml}}(T, S, I, Q),$$

and the value function (4.45) satisfies

$$\begin{aligned}\frac{\partial V^l}{\partial t} - rV^l + [\Theta^S(t) - \alpha_S \ln S]S\frac{\partial V^l}{\partial S} \\ + [\Theta^Z(t) - \alpha_Z \ln Z]Z\frac{\partial V^l}{\partial Z} + \frac{Z-M}{t-T_l}\frac{\partial V^l}{\partial M} \\ + \frac{1}{2}\sigma_S^2 S^2\frac{\partial^2 V^l}{\partial S^2} + \frac{1}{2}\sigma_Z^2 Z^2\frac{\partial^2 V^l}{\partial Z^2} + \sigma_S\sigma_Z\rho\frac{\partial^2 V^l}{\partial S\partial Z} \\ + \max_{q\in[0,\bar{q}]}\left\{q\left[(S-I) + \frac{\partial V^l}{\partial Q}\right]\right\} = 0 \quad (4.50)\end{aligned}$$

with the terminal condition (4.48). To avoid heavy notations, we have dropped the dependency of (t, S, I, Q) in V^{L-1}, the dependency of (t, S, Z, M, I, Q) in V^l and all corresponding partial derivatives.

Remark 4.7. The proof can be found in Theorem 2 and Theorem 8 of Edoli (2013). In Edoli (2013), the author also shows that (4.44) and (4.45) are viscosity solutions of (4.49) and (4.50), respectively. The theory of the viscosity solution is outside the scope of this thesis, we refer interested readers to Fleming and Soner (2006) and Basei, Cesaroni and Vargiolu (2014).

From (4.50) and (4.49), it can easily be seen that the optimal decision $q^*(t)$ at time t is

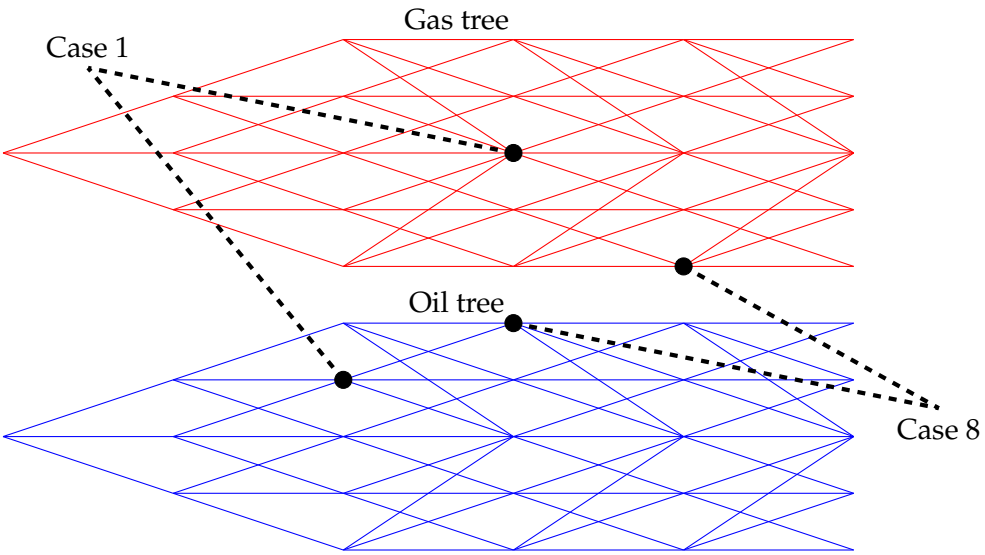
$$q^*(t) = \begin{cases} 0 & \text{if } S(t) - I(t) + \frac{\partial V^l}{\partial Q} \leq 0 \\ \bar{q} & \text{if } S(t) - I(t) + \frac{\partial V^l}{\partial Q} > 0 \end{cases}$$

That is, at time t, the optimal decision is either the daily minimum or the daily maximum.

4.9 First order consistency

Recall Remark 4.2, given a node $(n-1, s, z)$ where $s = -s_{\max}, -s_{\max}+1, \ldots, s_{\max}$ and $z = -z_{\max}, -z_{\max}+1, \ldots, z_{\max}$, the gas price and the oil price can be computed by (4.15). Let S^u, S^m and S^d be the gas price obtained by following the upper, middle and lower branches on the gas tree emanating from the node $(n-1, s, z)$, respectively. Similarly, let Z^u, Z^m and Z^d be the oil price obtained by following the upper, middle and lower branches on the

FIGURE 4.4: Illustration of Case 1 and Case 8



oil tree emanating from the node $(n-1, s, z)$, respectively. Then,

$$\begin{aligned} S^u &= S_{n,\mathfrak{g}(s,1)}, \quad & Z^u &= Z_{n,\mathfrak{h}(z,1)}, \\ S^m &= S_{n,\mathfrak{g}(s,0)}, \quad & Z^m &= Z_{n,\mathfrak{h}(z,0)}, \\ S^d &= S_{n,\mathfrak{g}(s,-1)}, \quad & Z^d &= Z_{n,\mathfrak{h}(z,-1)}. \end{aligned} \tag{4.51}$$

Recall (4.13) and (4.14), the values in (4.51) depend on the levels s and z. For the gas price, we have three possible outcomes:

1. $-s_{\max} < s < s_{\max}$, which means the node $(n-1, s, z)$ is not on the edges of the gas tree. In this case,

$$\begin{cases} S^u = S_{n,s+1} = e^{(s+1)\Delta X + a_n} \\ S^m = S_{n,s} = e^{s\Delta X + a_n} \\ S^d = S_{n,s-1} = e^{(s-1)\Delta X + a_n} \end{cases}$$

   In Assumption 4.1, we have assumed that $a_n = a(t_n)$. If we define the following notations:

$$S = S_{n-1,s}, \ X = s\Delta X, \text{ and } \ t = t_n,$$

we have

$$\begin{cases} S^u = S_{n,s+1} = e^{X+\Delta X+a(t)} \\ S^m = S_{n,s} = e^{X+a(t)} \\ S^d = S_{n,s-1} = e^{X-\Delta X+a(t)} \end{cases} \tag{4.52}$$

2. $s = s_{\max}$, which means the node $(n-1, s, z)$ is on the top edge of the gas tree. In this case,

$$\begin{cases} S^u = S_{n,s} = e^{X+a(t)} \\ S^m = S_{n,s-1} = e^{X-\Delta X+a(t)} \\ S^d = S_{n,s-2} = e^{X-2\Delta X+a(t)} \end{cases}$$

3. $s = -s_{\max}$, which means the node $(n-1, s, z)$ is on the bottom edge of the gas tree. In this case,

$$\begin{cases} S^u = S_{n,s+2} = e^{X+2\Delta X+a(t)} \\ S^m = S_{n,s+1} = e^{X+\Delta X+a(t)} \\ S^d = S_{n,s} = e^{X+a(t)} \end{cases}$$

Similarly, for the oil price, by defining

$$Z = Z_{n-1,z},\ Y = z\Delta Y,\ \text{and}\ \ t = t_n,$$

we also have three possible outcomes:

1. If $-z_{\max} < z < z_{\max}$, we have

$$\begin{cases} Z^u = Z_{n,z+1} = e^{Y+\Delta Y+b(t)} \\ Z^m = Z_{n,z} = e^{Y+b(t)} \\ Z^d = Z_{n,z-1} = e^{Y-\Delta Y+b(t)} \end{cases} \tag{4.53}$$

TABLE 4.6: Nine cases of node positions on a two-dimensional trinomial tree.

	Node positions
Case 1	$(n, \hat{s}, \hat{z})$
Case 2	$(n, \hat{s}, z_{\max})$
Case 3	$(n, \hat{s}, -z_{\max})$
Case 4	$(n, s_{\max}, \hat{z})$
Case 5	$(n, s_{\max}, z_{\max})$
Case 6	$(n, s_{\max}, -z_{\max})$
Case 7	$(n, -s_{\max}, \hat{z})$
Case 8	$(n, -s_{\max}, z_{\max})$
Case 9	$(n, -s_{\max}, -z_{\max})$

Note: $\hat{s}$ and $\hat{z}$ are integers such that $-s_{\max} < \hat{s} < s_{\max}$ and $-z_{\max} < \hat{z} < z_{\max}$.

2. If $z = z_{\max}$, we have

$$\begin{cases} Z^u = Z_{n,z} = e^{Y+b(t)} \\ Z^m = Z_{n,z-1} = e^{Y-\Delta Y+b(t)} \\ Z^d = Z_{n,z-2} = e^{Y-2\Delta Y+b(t)} \end{cases}$$

3. If $z = -z_{\max}$, we have

$$\begin{cases} Z^u = Z_{n,z} = e^{Y+2\Delta Y+b(t)} \\ Z^m = Z_{n,z-1} = e^{Y+\Delta Y+b(t)} \\ Z^d = Z_{n,z-2} = e^{Y+b(t)} \end{cases}$$

The combination of these possible outcomes of the gas price and the oil price gives us nine scenarios, as in Table 4.6. For the positions of these nodes in Table 4.6, Figure 4.4 shows examples of Case 1 and Case 8.

4.9.1 Proof of the first order consistency

In terms of the proof of the consistency, we give a detailed proof for Case 1 and the other cases can be proceeded similarly. First, rewrite (4.50) as follows:

$$\mathcal{L}V^l = 0, \tag{4.54}$$

where $\mathcal{L}$ is the operator such that

$$\begin{aligned}\mathcal{L}V^l =& \frac{\partial V^l}{\partial t} - rV^l + [\Theta^S(t) - \alpha_S \ln S]S\frac{\partial V^l}{\partial S} \\ &+ [\Theta^Z(t) - \alpha_Z \ln Z]Z\frac{\partial V^l}{\partial Z} + \frac{Z-M}{t-T_l}\frac{\partial V^l}{\partial M} \\ &+ \frac{1}{2}\sigma_S^2 S^2\frac{\partial^2 V^l}{\partial S^2} + \frac{1}{2}\sigma_Z^2 Z^2\frac{\partial^2 V^l}{\partial Z^2} + \sigma_S\sigma_Z\rho\frac{\partial^2 V^l}{\partial S\partial Z} \\ &+ \max_{q\in[0,\bar{q}]}\left\{q\left[(S-I) + \frac{\partial V^l}{\partial Q}\right]\right\},\end{aligned}$$

where

$$\begin{aligned}\Theta^S(t) &= \frac{\partial a(t)}{\partial t} + \alpha_S a(t) + \frac{1}{2}\sigma_S^2, \\ \Theta^Z(t) &= \frac{\partial b(t)}{\partial t} + \alpha_Z b(t) + \frac{1}{2}\sigma_Z^2.\end{aligned}$$

Recall our algorithm in Section 4.6. The combination of equations (4.28) and (4.29) gives that, for $t_n \in (T_l, T_{l+1}]$, $l = 1, 2, \ldots, L-2$,

$$\begin{aligned}V^l(S_{n,s}, Z_{n,z}, M_{n,m}, I_{n,i}, Q) = \max_{q\in[0,\bar{q}]}\Bigg\{& q\cdot(S_{n,s} - I_{n,i}) + \\ &\sum_{b,c=-1}^{1} p_{s,z,b,c}e^{-r\Delta t}\cdot \\ &V^l(S_{n+1,\mathfrak{g}(s,b)}, Z_{n+1,\mathfrak{h}(z,c)}, \mathcal{M}_{n,m,c}, I_{n,i}, Q+q)\Bigg\}.\end{aligned} \tag{4.55}$$

Until indicated otherwise, we ignore the effect of interpolation (4.32). Now, we show that the consistency of the tree algorithm (4.55). Denote $\mathcal{D} = (0,\infty)\times(0,\infty)\times(0,\infty)\times(0,\infty)\times[0,\bar{q}T]$. We need to show that, for sufficiently smooth

function $U(t,S,Z,M,I,Q)$ and $\mathbf{D}_0=(t_0,S_0,Z_0,M_0,I_0,Q_0)\in(T_l,T_{l+1})\times\mathcal{D}$,

$$\lim_{\substack{\Delta t\to 0\\(t,S,Z,M,I,Q)\to\mathbf{D}_0}}\frac{1}{\Delta t}(U-F_{\Delta t}U)(t-\Delta t,S,Z,M,I,Q)=-\mathcal{L}U\Big|_{\mathbf{D}_0}, \tag{4.56}$$

where $F_{\Delta t}U(t,S,Z,M,I,Q)$ satisfies

$$\begin{aligned}&F_{\Delta t}U(t-\Delta t,S,Z,M,I,Q)\\&=\max_{q\in[0,\bar{q}]}\Big\{q\cdot\Delta t\cdot\Big(S-I\Big)+e^{-r\Delta t}\mathcal{U}(t,S,Z,M,I,Q)\Big\}.\end{aligned} \tag{4.57}$$

Depending on the positions of the nodes on the two-dimensional trinomial tree, $\mathcal{U}(t,S,Z,M,I,Q)$ in (4.57) have different forms. In Case 1, we have

$$\begin{aligned}\mathcal{U}(t,S,Z,M,I,Q)=&P_{uu}\cdot U\big(t,S^u,Z^u,M^u,I,Q+q\cdot\Delta t\big)+\\&P_{um}\cdot U\big(t,S^u,Z^m,M^m,I,Q+q\cdot\Delta t\big)+\\&P_{ud}\cdot U\big(t,S^u,Z^d,M^d,I,Q+q\cdot\Delta t\big)+\\&P_{mu}\cdot U\big(t,S^m,Z^u,M^u,I,Q+q\cdot\Delta t\big)+\\&P_{mm}\cdot U\big(t,S^m,Z^m,M^m,I,Q+q\cdot\Delta t\big)+\\&P_{md}\cdot U\big(t,S^m,Z^d,M^d,I,Q+q\cdot\Delta t\big)+\\&P_{du}\cdot U\big(t,S^d,Z^u,M^u,I,Q+q\cdot\Delta t\big)+\\&P_{dm}\cdot U\big(t,S^d,Z^m,M^m,I,Q+q\cdot\Delta t\big)+\\&P_{dd}\cdot U\big(t,S^d,Z^d,M^d,I,Q+q\cdot\Delta t\big),\end{aligned} \tag{4.58}$$

where S^u, S^m, S^d, Z^u, Z^m and Z^d are given by (4.52) and (4.53). M^u, M^m and M^d are the values of the running average if the oil price moves from Z to Z^u, Z^m and Z^d, respectively. In the same sprit as (4.18), together with (4.53), we have

$$\begin{cases}M^u&=\dfrac{(t-\Delta t-T_l)\cdot M+Z^u\Delta t}{t-T_l},\\M^m&=\dfrac{(t-\Delta t-T_l)\cdot M+Z^m\Delta t}{t-T_l},\\M^d&=\dfrac{(t-\Delta t-T_l)\cdot M+Z^d\Delta t}{t-T_l}.\end{cases} \tag{4.59}$$

Recall the construction of a two-dimensional trinomial tree in Section 2.2, P_{uu}, P_{um}, P_{ud}, P_{mu}, P_{mm}, P_{md}, P_{du}, P_{dm} and P_{dd} are probabilities associated

with the nine possible movements on the two-dimensional trinomial tree. These probabilities are calculated through (2.27), (2.28) and (2.29). Denote the probabilities associated with the upper, middle and lower branches on the gas tree by p_u, p_m and p_d, respectively. Similarly, denote the probabilities associated with the upper, middle and lower branches on the oil tree by q_u, q_m and q_d, respectively. p_u, p_m, p_d, q_u, q_m and q_d can be calculated by using (2.26) (also see Remark 2.8).

By (2.26), we have

$$\begin{cases} p_u = \frac{1}{6} + \mathcal{O}(\Delta t), \\ p_m = \frac{2}{3} + \mathcal{O}(\Delta t^2), \\ p_d = \frac{1}{6} + \mathcal{O}(\Delta t), \end{cases} \qquad \begin{cases} q_u = \frac{1}{6} + \mathcal{O}(\Delta t), \\ q_m = \frac{2}{3} + \mathcal{O}(\Delta t^2), \\ q_d = \frac{1}{6} + \mathcal{O}(\Delta t). \end{cases} \tag{4.60}$$

In Section 2.2, these probabilities are obtained matching the theoretical conditional expectation and variance (see (2.21) and (2.22)). In addition, recall Remark 4.2, ΔX and ΔY are the space steps on the gas fundamental tree and the oil fundamental tree, respectively. The value of ΔX and ΔY are given by

$$\Delta X = \sigma_S \sqrt{3\Delta t} \text{ and } \Delta Y = \sigma_Z \sqrt{3\Delta t}, \tag{4.61}$$

respectively. Therefore, using (2.21), (2.22) and (4.61), we have the following equations:

$$\begin{cases} p_u \cdot \Delta X + p_m \cdot 0 + p_d \cdot (-\Delta X) = -\alpha_S X \Delta t \\ q_u \cdot \Delta Y + q_m \cdot 0 + q_d \cdot (-\Delta Y) = -\alpha_Z Y \Delta t \\ p_u \cdot (\Delta X)^2 + p_m \cdot 0 + p_d \cdot (-\Delta X)^2 = \sigma_S^2 \Delta t + \mathcal{O}(\Delta t^2) \\ q_u \cdot (\Delta Y)^2 + q_m \cdot 0 + q_d \cdot (-\Delta Y)^2 = \sigma_Z^2 \Delta t + \mathcal{O}(\Delta t^2) \end{cases}. \tag{4.62}$$

Again, from the tree building procedures in Section 2.2, we have

$$P_{uu} + P_{um} + P_{ud} + P_{mu} + P_{mm} + P_{md} + P_{du} + P_{dm} + P_{dd} = 1. \tag{4.63}$$

Using (4.61) and the Taylor expansions, we derive some equations in (4.64) which are intensively used in the rest of this section:

$$\begin{cases} e^{\Delta X} = 1 + \Delta X + \frac{1}{2}\Delta X^2 + \frac{1}{6}\Delta X^3 + \mathcal{O}(\Delta t^2), \\ e^{-\Delta X} = 1 - \Delta X + \frac{1}{2}\Delta X^2 - \frac{1}{6}\Delta X^3 + \mathcal{O}(\Delta t^2), \\ e^{a(t)} = e^{a(t-\Delta t)} + \Delta t e^{a(t-\Delta t)} \dfrac{\partial a(t-\Delta t)}{\partial t} + \mathcal{O}(\Delta t^2), \\ a(t-\Delta t) = a(t) + \mathcal{O}(\Delta t), \\ \dfrac{\partial a(t-\Delta t)}{t} = \dfrac{\partial a(t)}{\partial t} + \mathcal{O}(\Delta t). \end{cases} \tag{4.64}$$

From (4.64), we can have similar results for $e^{\Delta Y}$, $e^{-\Delta Y}$, $e^{b(t)}$, $b(t-\Delta t)$ and $\frac{\partial b(t-\Delta t)}{t}$. Define the following notations:

$$\begin{array}{ccc} \Delta S^u = (S^u - S), & \Delta S^m = (S^m - S), & \Delta S^d = (S^d - S), \\ \Delta Z^u = (Z^u - Z), & \Delta Z^m = (Z^m - Z), & \Delta Z^d = (S^d - Z), \\ \Delta M^u = (M^u - M), & \Delta M^m = (M^m - M), & \Delta M^d = (M^d - M). \end{array}$$

Using the equalities in (4.64), we can further derive the following equalities using the Taylor expansions:

$$\begin{cases} \Delta S^u = \Big[\Delta X + \frac{1}{2}\Delta X^2 + \frac{1}{6}\Delta X^3 + \Delta X \Delta t \frac{\partial a(t)}{\partial t} + \Delta t \frac{\partial a(t)}{\partial t}\Big] S + \mathcal{O}(\Delta t^2), \\ \Delta S^m = \dfrac{\partial a(t)}{\partial t} S \Delta t + \mathcal{O}(\Delta t^2), \\ \Delta S^d = \Big[- \Delta X + \frac{1}{2}\Delta X^2 - \frac{1}{6}\Delta X^3 - \Delta X \Delta t \frac{\partial a(t)}{\partial t} + \Delta t \frac{\partial a(t)}{\partial t}\Big] S + \mathcal{O}(\Delta t^2), \\ \Delta Z^u = \Big[\Delta Y + \frac{1}{2}\Delta Y^2 + \frac{1}{6}\Delta Y^3 + \Delta Y \Delta t \frac{\partial b(t)}{\partial t} + \Delta t \frac{\partial b(t)}{\partial t}\Big] Z + \mathcal{O}(\Delta t^2), \\ \Delta Z^m = \dfrac{\partial b(t)}{\partial t} Z \Delta t + \mathcal{O}(\Delta t^2), \\ \Delta Z^d = \Big[- \Delta Y + \frac{1}{2}\Delta Y^2 - \frac{1}{6}\Delta Y^3 - \Delta Y \Delta t \frac{\partial b(t)}{\partial t} + \Delta t \frac{\partial b(t)}{\partial t}\Big] Z + \mathcal{O}(\Delta t^2), \\ \Delta M^u = \dfrac{Z(1+\Delta Y) - M}{t - T_l} \cdot \Delta t + \mathcal{O}(\Delta t^2), \\ \Delta M^m = \dfrac{Z - M}{t - T_l} \cdot \Delta t + \mathcal{O}(\Delta t^2), \\ \Delta M^d = \dfrac{Z(1-\Delta Y) - M}{t - T_l} \cdot \Delta t + \mathcal{O}(\Delta t^2). \end{cases} \tag{4.65}$$

Here, we give explanations on how we derive ΔS^u, the rest can be proceeded similarly:

$$\begin{aligned}
\Delta S^u =& e^{X+\Delta X+a(t)} - e^{X+a(t-\Delta t)} \\
=& e^{X+\Delta X}\Big[e^{a(t-\Delta t)} + \Delta t\frac{\partial a(t-\Delta t)}{\partial t}e^{a(t-\Delta t)}\Big] - e^{X+a(t-\Delta t)} + \mathcal{O}(\Delta t^2) \\
=& \Big[e^{\Delta X} + \Delta t e^{\Delta X}\frac{\partial a(t-\Delta t)}{\partial t} - 1\Big]S + \mathcal{O}(\Delta t^2) \\
=& \Big[\Delta X + \frac{1}{2}\Delta X^2 + \frac{1}{6}\Delta X^3 + \big[1+\Delta X\big]\Delta t\frac{\partial a(t-\Delta t)}{\partial t}\Big]S + \mathcal{O}(\Delta t^2) \\
=& \Big[\Delta X + \frac{1}{2}\Delta X^2 + \frac{1}{6}\Delta X^3 + \Delta X\Delta t\frac{\partial a(t)}{\partial t} + \Delta t\frac{\partial a(t)}{\partial t}\Big]S + \mathcal{O}(\Delta t^2).
\end{aligned}$$

Using (4.65), we have the following results:

$$\begin{cases}
\Delta t\Delta M^u = \mathcal{O}(\Delta t^2),\ \Delta t\Delta M^m = \mathcal{O}(\Delta t^2),\ \Delta t\Delta M^d = \mathcal{O}(\Delta t^2), \\
(\Delta M^u)^2 = \mathcal{O}(\Delta t^2),\ (\Delta M^m)^2 = \mathcal{O}(\Delta t^2),\ (\Delta M^d)^2 = \mathcal{O}(\Delta t^2), \\
(\Delta S^u)^2 = \Delta X^2 + \Delta X^3 + \mathcal{O}(\Delta t^2), \\
(\Delta S^m)^2 = \mathcal{O}(\Delta t^2), \\
(\Delta S^d)^2 = \Delta X^2 - \Delta X^3 + \mathcal{O}(\Delta t^2), \\
(\Delta Z^u)^2 = \Delta Y^2 + \Delta Y^3 + \mathcal{O}(\Delta t^2), \\
(\Delta Z^m)^2 = \mathcal{O}(\Delta t^2), \\
(\Delta Z^d)^2 = \Delta Y^2 - \Delta Y^3 + \mathcal{O}(\Delta t^2), \\
(\Delta S^u)^3 = \Delta X^3 + \mathcal{O}(\Delta t^2),\ (\Delta S^m)^3 = \mathcal{O}(\Delta t^3),\ (\Delta S^d)^3 = -\Delta X^3 + \mathcal{O}(\Delta t^2), \\
(\Delta Z^u)^3 = \Delta Y^3 + \mathcal{O}(\Delta t^2),\ (\Delta Z^m)^3 = \mathcal{O}(\Delta t^3),\ (\Delta Z^d)^3 = -\Delta Y^3 + \mathcal{O}(\Delta t^2).
\end{cases} \tag{4.66}$$

(4.65) together with (4.60), (4.62) and (4.66), one can show that the following equalities hold:

$$\begin{cases}
q_u\cdot\Delta M^u + q_m\cdot\Delta M^m + q_d\cdot\Delta M^d = \dfrac{Z-M}{t-T_l}\cdot\Delta t + \mathcal{O}(\Delta t^2), \\
p_u\cdot\Delta S^u + p_m\cdot\Delta S^m + p_d\cdot\Delta S^d = \big[\Theta^S(t) - \alpha_S\ln S\big]S\Delta t + \mathcal{O}(\Delta t^2), \\
q_u\cdot\Delta Z^u + q_m\cdot\Delta Z^u + q_d\cdot\Delta Z^u = \big[\Theta^Z(t) - \alpha_Z\ln Z\big]Z\Delta t + \mathcal{O}(\Delta t^2), \\
p_u(\Delta S^u)^2 + p_m(\Delta S^m)^2 + p_d(\Delta S^d)^2 = \sigma_S^2S^2\Delta t + \mathcal{O}(\Delta t^2), \\
q_u(\Delta Z^u)^2 + q_m(\Delta Z^m)^2 + q_d(\Delta Z^d)^2 = \sigma_Z^2Z^2\Delta t + \mathcal{O}(\Delta t^2), \\
p_u(\Delta S^u)^3 + p_m(\Delta S^m)^3 + p_d(\Delta S^d)^3 = \mathcal{O}(\Delta t^2), \\
q_u(\Delta Z^u)^3 + q_m(\Delta Z^m)^3 + q_d(\Delta Z^d)^3 = \mathcal{O}(\Delta t^2).
\end{cases} \tag{4.67}$$

Here, we give explanations on how we derive $p_u \cdot \Delta S^u + p_m \cdot \Delta S^m + p_d \cdot \Delta S^d$, the rest can be proceeded similarly:

$$
\begin{aligned}
& p_u \cdot \Delta S^u + p_m \cdot \Delta S^m + p_d \cdot \Delta S^d \\
=& \Big[-\alpha_S X + \frac{1}{2}\sigma_S^2 + \frac{\partial a(t)}{\partial t} \Big] S\Delta t + \mathcal{O}(\Delta t^2) \\
=& \Big[\alpha_S a(t-\Delta t) - \alpha_S \ln S + \frac{1}{2}\sigma_S^2 + \frac{\partial a(t)}{\partial t} \Big] S\Delta t + \mathcal{O}(\Delta t^2) \\
=& \Big[\alpha_S a(t) - \alpha_S \ln S + \frac{1}{2}\sigma_S^2 + \frac{\partial a(t)}{\partial t} \Big] S\Delta t + \mathcal{O}(\Delta t^2) \\
=& \big[\Theta^S(t) - \alpha_S \ln S \big] S\Delta t + \mathcal{O}(\Delta t^2).
\end{aligned}
$$

Starting from (4.65), using (4.61) and (4.64), we have the following equalities:

$$
\begin{cases}
\Delta Z^u \Delta M^u =& \dfrac{\Delta Y \cdot Z - M}{t - T_l} \Delta t + \mathcal{O}(\Delta t^2), \\
\Delta Z^m \Delta M^m =& \mathcal{O}(\Delta t^2), \\
\Delta Z^d \Delta M^d =& \dfrac{-\Delta Y \cdot Z + M}{t - T_l} \Delta t + \mathcal{O}(\Delta t^2), \\
\Delta S^u \Delta Z^u =& \Big[\Delta X \Delta Y + \frac{1}{2}\Delta X \Delta Y^2 + \frac{1}{2}\Delta Y \Delta X^2 \\
& + \Delta X \Delta t \frac{\partial b(t)}{\partial t} + \Delta Y \Delta t \frac{\partial a(t)}{\partial t} \Big] SZ + \mathcal{O}(\Delta t^2), \\
\Delta S^u \Delta Z^m =& \Big[\Delta X \Delta t \frac{\partial b(t)}{\partial t} \Big] SZ + \mathcal{O}(\Delta t^2), \\
\Delta S^u \Delta Z^d =& \Big[-\Delta X \Delta Y + \frac{1}{2}\Delta X \Delta Y^2 - \frac{1}{2}\Delta Y \Delta X^2 \\
& + \Delta X \Delta t \frac{\partial b(t)}{\partial t} - \Delta Y \Delta t \frac{\partial a(t)}{\partial t} \Big] SZ + \mathcal{O}(\Delta t^2), \\
\Delta S^m \Delta Z^u =& \Big[\Delta Y \Delta t \frac{\partial a(t)}{\partial t} \Big] SZ + \mathcal{O}(\Delta t^2), \\
\Delta S^m \Delta Z^m =& \mathcal{O}(\Delta t^2), \\
\Delta S^m \Delta Z^d =& \Big[-\Delta Y \Delta t \frac{\partial a(t)}{\partial t} \Big] SZ + \mathcal{O}(\Delta t^2), \\
\Delta S^d \Delta Z^u =& \Big[-\Delta X \Delta Y - \frac{1}{2}\Delta X \Delta Y^2 + \frac{1}{2}\Delta Y \Delta X^2 \\
& - \Delta X \Delta t \frac{\partial b(t)}{\partial t} + \Delta Y \Delta t \frac{\partial a(t)}{\partial t} \Big] SZ + \mathcal{O}(\Delta t^2), \\
\Delta S^d \Delta Z^m =& \Big[-\Delta X \Delta t \frac{\partial b(t)}{\partial t} \Big] SZ + \mathcal{O}(\Delta t^2), \\
\Delta S^d \Delta Z^d =& \Big[\Delta X \Delta Y - \frac{1}{2}\Delta X \Delta Y^2 - \frac{1}{2}\Delta Y \Delta X^2 \\
& - \Delta X \Delta t \frac{\partial b(t)}{\partial t} - \Delta Y \Delta t \frac{\partial a(t)}{\partial t} \Big] SZ + \mathcal{O}(\Delta t^2).
\end{cases}
\tag{4.68}
$$

Using (4.68) together with (4.65), we have

$$\Delta S \Delta Z \Delta M = \mathcal{O}(\Delta t^2) \tag{4.69}$$

for any $\Delta S \in \{\Delta S^u, \Delta S^m, \Delta S^d\}$, $\Delta Z \in \{\Delta Z^u, \Delta Z^m, \Delta Z^d\}$ and $\Delta M \in \{\Delta M^u, \Delta M^m, \Delta M^d\}$.

Recall ε in (2.28) and (2.29). $\varepsilon = \frac{\rho}{36}$ if $\rho \geq 0$ and $\varepsilon = -\frac{\rho}{36}$ if $\rho < 0$. Using the probability matrix (2.27), (2.28) and (2.29), we have

$$\begin{cases} \Big(P_{uu} \cdot \Delta S^u + P_{mu} \cdot \Delta S^m + P_{du} \cdot \Delta S^d\Big)\Delta M^u \\ = \begin{cases} -6\varepsilon \Delta X \Delta M^u + \mathcal{O}(\Delta t^2), & \rho \geq 0, \\ 6\varepsilon \Delta X \Delta M^u + \mathcal{O}(\Delta t^2), & \rho < 0, \end{cases} \\ \Big(P_{um} \cdot \Delta S^u + P_{mm} \cdot \Delta S^m + P_{dm} \cdot \Delta S^d\Big)\Delta M^m \\ = \mathcal{O}(\Delta t^2), \\ \Big(P_{ud} \cdot \Delta S^u + P_{md} \cdot \Delta S^m + P_{dd} \cdot \Delta S^d\Big)\Delta M^d \\ = \begin{cases} -6\varepsilon \Delta X \Delta M^d + \mathcal{O}(\Delta t^2), & \rho \geq 0, \\ 6\varepsilon \Delta X \Delta M^d + \mathcal{O}(\Delta t^2), & \rho < 0, \end{cases} \end{cases}$$

which gives

$$\begin{aligned} \big(P_{uu} \cdot \Delta S^u + P_{mu} \cdot \Delta S^m + P_{du} \cdot \Delta S^d\big)\Delta M^u + \big(P_{um} \cdot \Delta S^u + \\ P_{mm} \cdot \Delta S^m + P_{dm} \cdot \Delta S^d\big)\Delta M^m + \big(P_{ud} \cdot \Delta S^u + \\ P_{md} \cdot \Delta S^m + P_{dd} \cdot \Delta S^d\big)\Delta M^d = \mathcal{O}(\Delta t^2). \end{aligned} \tag{4.70}$$

(4.68) together with (4.60), (4.62), (2.27), (2.28) and (2.29), one can derive

$$\begin{aligned} & P_{uu} \cdot \Delta S^u \cdot \Delta Z^u + P_{mu} \cdot \Delta S^m \cdot \Delta Z^u + P_{du} \cdot \Delta S^d \cdot \Delta Z^u \\ & + P_{um} \cdot \Delta S^u \cdot \Delta Z^m + P_{mm} \cdot \Delta S^m \cdot \Delta Z^m + P_{dm} \cdot \Delta S^d \cdot \Delta Z^m \\ & + P_{ud} \cdot \Delta S^u \cdot \Delta Z^d + P_{md} \cdot \Delta S^m \cdot \Delta Z^d + P_{dd} \cdot \Delta S^d \cdot \Delta Z^d \\ = & \begin{cases} 12\varepsilon \cdot \Delta X \cdot \Delta Y + \mathcal{O}(\Delta t^2), & \rho \geq 0 \\ -12\varepsilon \cdot \Delta X \cdot \Delta Y + \mathcal{O}(\Delta t^2), & \rho < 0 \end{cases} \\ = & \rho \sigma_S \sigma_Z \Delta t + \mathcal{O}(\Delta t^2). \end{aligned} \tag{4.71}$$

(4.68) together with (4.60) and (4.62), we can also have

$$q_u \cdot \Delta Z^u \Delta M^u + q_m \cdot \Delta Z^m \Delta M^m + q_d \cdot \Delta Z^d \Delta M^d = \mathcal{O}(\Delta t^2). \tag{4.72}$$

Now, we perform the Taylor expansions of (4.58) at (t, S, Z, M, I, Q). Together with the result in (4.69), we have

$$\begin{aligned}
&\mathcal{U}(t, S, Z, M, I, Q) \\
=&U\big(t, S, Z, M, I, Q\big) + \mathbf{A}_1 \frac{\partial U}{\partial S} + \mathbf{A}_2 \frac{\partial U}{\partial Z} + \mathbf{A}_3 \frac{\partial U}{\partial M} + \mathbf{A}_4 \frac{\partial^2 U}{\partial S^2} + \mathbf{A}_5 \frac{\partial^2 U}{\partial Z^2} \\
&+ \mathbf{A}_6 \frac{\partial^2 U}{\partial S \partial Z} + \mathbf{A}_7 \frac{\partial U}{\partial Q} + \mathbf{A}_8 \frac{\partial^2 U}{\partial S \partial M} + \mathbf{A}_9 \frac{\partial^2 U}{\partial Z \partial M} + \mathbf{A}_{10} \frac{\partial^2 U}{\partial S \partial Q} + \mathbf{A}_{11} \frac{\partial^2 U}{\partial Z \partial Q} \\
&+ \mathbf{A}_{12} \frac{\partial^3 U}{\partial S^3} + \mathbf{A}_{13} \frac{\partial^3 U}{\partial Z^3} + \mathcal{O}(\Delta t^2),
\end{aligned} \tag{4.73}$$

where $\mathbf{A}_1$-$\mathbf{A}_{13}$ are the functions in front of these derivatives obtained through the Taylor expansions.

After using the Taylor expansions, $\mathbf{A}_1$ is given by

$$\mathbf{A}_1 = (P_{uu}+P_{um}+P_{ud})\Delta S^u + (P_{mu}+P_{mm}+P_{md})\Delta S^m + (P_{du}+P_{dm}+P_{dd})\Delta S^d.$$

Using (2.27), (2.28) and (2.29), we have

$$\mathbf{A}_1 = p_u \Delta S^u + p_m \Delta S^m + p_d \Delta S^d.$$

From (4.67), we know

$$\mathbf{A}_1 = \big[\Theta^S(t) - \alpha_S \ln S\big] S \Delta t + \mathcal{O}(\Delta t^2). \tag{4.74}$$

Similarly, we can get

$$\mathbf{A}_2 = \big[\Theta^Z(t) - \alpha_Z \ln Z\big] Z \Delta t + \mathcal{O}(\Delta t^2).$$

For $\mathbf{A}_3$, after using the Taylor expansions, we have

$$\begin{aligned}
\mathbf{A}_3 =&(P_{uu} + P_{mu} + P_{du})\Delta M^u + (P_{um} + P_{mm} + P_{dm})\Delta M^m \\
&+ (P_{ud} + P_{md} + P_{dd})\Delta M^d \\
=&q_u \Delta M^u + q_m \Delta M^m + q_d \Delta M^d.
\end{aligned}$$

By (4.67), we have

$$\mathbf{A}_3 = \frac{Z - M}{t - T_l} \cdot \Delta t + \mathcal{O}(\Delta t^2).$$

For $\mathbf{A}_4$, we have

$$\begin{aligned}\mathbf{A}_4 =& \frac{1}{2}\Big[(P_{uu} + P_{um} + P_{ud})(\Delta S^u)^2 + (P_{mu} + P_{mm} + P_{md})(\Delta S^m)^2 \\ &+ (P_{du} + P_{dm} + P_{dd})(\Delta S^d)^2\Big] \\ =& \frac{1}{2}\Big[p_u(\Delta S^u)^2 + p_m(\Delta S^m)^2 + p_d(\Delta S^d)^2\Big].\end{aligned}$$

Together with (4.67), we have

$$\mathbf{A}_4 = \frac{1}{2}\sigma_S^2 S^2 \Delta t + \mathcal{O}(\Delta t^2).$$

Similarly, we have

$$\mathbf{A}_5 = \frac{1}{2}\sigma_Z^2 Z^2 \Delta t + \mathcal{O}(\Delta t^2).$$

For $\mathbf{A}_6$, we have

$$\begin{aligned}\mathbf{A}_6 =& P_{uu} \cdot \Delta S^u \cdot \Delta Z^u + P_{mu} \cdot \Delta S^m \cdot \Delta Z^u + P_{du} \cdot \Delta S^d \cdot \Delta Z^u \\ &+ P_{um} \cdot \Delta S^u \cdot \Delta Z^m + P_{mm} \cdot \Delta S^m \cdot \Delta Z^m + P_{dm} \cdot \Delta S^d \cdot \Delta Z^m \\ &+ P_{ud} \cdot \Delta S^u \cdot \Delta Z^d + P_{md} \cdot \Delta S^m \cdot \Delta Z^d + P_{dd} \cdot \Delta S^d \cdot \Delta Z^d.\end{aligned}$$

From (4.71), we have

$$\mathbf{A}_6 = \rho\sigma_S\sigma_Z \Delta t + \mathcal{O}(\Delta t^2).$$

For $\mathbf{A}_7$, we have

$$\mathbf{A}_7 = q \cdot \Delta t\big(P_{uu} + P_{um} + P_{ud} + P_{mu} + P_{mm} + P_{md} + P_{du} + P_{dm} + P_{dd}\big).$$

From (4.63), we have

$$\mathbf{A}_7 = q \cdot \Delta t.$$

For $\mathbf{A}_8$, we have

$$\begin{aligned}\mathbf{A}_8 =& P_{uu}\cdot\Delta S^u\cdot\Delta M^u + P_{mu}\cdot\Delta S^m\cdot\Delta M^u + P_{du}\cdot\Delta S^d\cdot\Delta M^u\\ &+ P_{um}\cdot\Delta S^u\cdot\Delta M^m + P_{mm}\cdot\Delta S^m\cdot\Delta M^m + P_{dm}\cdot\Delta S^d\cdot\Delta M^m\\ &+ P_{ud}\cdot\Delta S^u\cdot\Delta M^d + P_{md}\cdot\Delta S^m\cdot\Delta M^d + P_{dd}\cdot\Delta S^d\cdot\Delta M^d.\end{aligned}$$

From (4.70), we have

$$\mathbf{A}_8 = \mathcal{O}(\Delta t^2).$$

For $\mathbf{A}_9$, we have

$$\begin{aligned}\mathbf{A}_9 =& (P_{uu}+P_{mu}+P_{du})\Delta Z^u\cdot\Delta M^u + (P_{um}+P_{mm}+P_{dm})\Delta Z^m\cdot\Delta M^m\\ &+ (P_{ud}+P_{md}+P_{dd})\Delta Z^d\cdot\Delta M^d\\ =& q_u\cdot\Delta Z^u\cdot\Delta M^u + q_m\cdot\Delta Z^m\cdot\Delta M^m + q_d\cdot\Delta Z^d\cdot\Delta M^d.\end{aligned}$$

From (4.72), we have

$$\mathbf{A}_9 = \mathcal{O}(\Delta t^2).$$

For $\mathbf{A}_{10}$,

$$\begin{aligned}\mathbf{A}_{10} =& \Big[(P_{uu}+P_{um}+P_{ud})\Delta S^u + (P_{mu}+P_{mm}+P_{md})\Delta S^m\\ &+ (P_{du}+P_{dm}+P_{dd})\Delta S^d\Big]\cdot q\cdot\Delta t\\ =& \mathbf{A}_1\cdot q\cdot\Delta t.\end{aligned}$$

Since $\mathbf{A}_1 = \big[\Theta^S(t) - \alpha_S \ln S\big]S\Delta t + \mathcal{O}(\Delta t^2)$ (see (4.74)), it follows that

$$\mathbf{A}_{10} = \mathcal{O}(\Delta t^2).$$

Similarly, we have

$$\mathbf{A}_{11} = \mathcal{O}(\Delta t^2).$$

For $\mathbf{A}_{12}$,

$$\begin{aligned}\mathbf{A}_{12} =&\frac{1}{6}\Big[(P_{uu}+P_{um}+P_{ud})(\Delta S^u)^3+(P_{mu}+P_{mm}+P_{md})(\Delta S^m)^3\\&+(P_{du}+P_{dm}+P_{dd})(\Delta S^d)^3\Big]\\=&\frac{1}{6}\Big[p_u(\Delta S^u)^3+p_m(\Delta S^m)^3+p_d(\Delta S^d)^3\Big].\end{aligned}$$

From (4.67), we have

$$\mathbf{A}_{12}=\mathcal{O}(\Delta t^2).$$

Similarly, we have

$$\mathbf{A}_{13}=\mathcal{O}(\Delta t^2).$$

Put $\mathbf{A}_1$-$\mathbf{A}_{13}$ into (4.73), we obtain

$$\begin{aligned}&\mathcal{U}(t,S,Z,M,I,Q)\\=&U\big(t,S,Z,M,I,Q\big)+\Big\{\big[\Theta^S(t)-\alpha_S\ln S\big]S\frac{\partial U}{\partial S}+\big[\Theta^Z(t)-\alpha_Z\ln Z\big]Z\frac{\partial U}{\partial Z}\\&+\frac{Z-M}{t-T_l}\frac{\partial U}{\partial M}+q\cdot\frac{\partial U}{\partial Q}+\frac{1}{2}\sigma_S^2S^2\frac{\partial^2U}{\partial S^2}+\frac{1}{2}\sigma_Z^2Z^2\frac{\partial^2U}{\partial Z^2}\\&+\rho\sigma_S\sigma_ZSZ\frac{\partial^2U}{\partial S\partial Z}\Big\}\Delta t+\mathcal{O}(\Delta t^2).\end{aligned}$$

Since

$$e^{-r\Delta t}=1-r\Delta t+\mathcal{O}(\Delta t^2),$$

it follows that

$$\begin{aligned}&e^{-r\Delta t}\mathcal{U}(t,S,Z,M,I,Q)\\=&U\big(t,S,Z,M,I,Q\big)+\Big\{\big[\Theta^S(t)-\alpha_S\ln S\big]S\frac{\partial U}{\partial S}+\big[\Theta^Z(t)-\alpha_Z\ln Z\big]Z\frac{\partial U}{\partial Z}\\&+\frac{Z-M}{t-T_l}\frac{\partial U}{\partial M}+\frac{1}{2}\sigma_S^2S^2\frac{\partial^2U}{\partial S^2}+\frac{1}{2}\sigma_Z^2Z^2\frac{\partial^2U}{\partial Z^2}\\&+\rho\sigma_S\sigma_ZSZ\frac{\partial^2U}{\partial S\partial Z}-rU\big(t,S,Z,M,I,Q\big)\Big\}\Delta t+\mathcal{O}(\Delta t^2).\end{aligned}\tag{4.75}$$

Put (4.75) back into (4.57), we have the following result:

$$\begin{aligned}
&F_{\Delta t}U(t-\Delta t,S,Z,M,I,Q)\\
=&U\big(t,S,Z,M,I,Q\big)+\Big\{\big[\Theta^S(t)-\alpha_S\ln S\big]S\frac{\partial U}{\partial S}+\big[\Theta^Z(t)-\alpha_Z\ln Z\big]Z\frac{\partial U}{\partial Z}\\
&+\frac{Z-M}{t-T_l}\frac{\partial U}{\partial M}+q\cdot\frac{\partial U}{\partial Q}+\frac{1}{2}\sigma_S^2S^2\frac{\partial^2U}{\partial S^2}+\frac{1}{2}\sigma_Z^2Z^2\frac{\partial^2U}{\partial Z^2}+\rho\sigma_S\sigma_ZSZ\frac{\partial^2U}{\partial S\partial Z}\\
&-rU\big(t,S,Z,M,I,Q\big)\Big\}\Delta t+\max_{q\in[0,\bar{q}]}\left\{q\cdot\Delta t\cdot\Big(S-I\Big)+q\cdot\Delta t\frac{\partial U}{\partial Q}\right\}\\
&+\mathcal{O}(\Delta t^2) \qquad (4.76)
\end{aligned}$$

Using Taylor expansions of $U\big(t-\Delta t,S,Z,M,I,Q\big)$ on (t,S,Z,M,I,Q), we have

$$U\big(t-\Delta t,S,Z,M,I,Q\big)=U\big(t,S,Z,M,I,Q\big)-\Delta t\frac{\partial U}{\partial t}+\mathcal{O}(\Delta t^2). \qquad (4.77)$$

Using (4.76) and (4.77), we have

$$\begin{aligned}
&\frac{1}{\Delta t}(U-F_{\Delta t}U)(t-\Delta t,S,Z,M,I,Q)\\
=&-\frac{\partial U}{\partial t}+rU\big(t,S,Z,M,I,Q\big)-\big[\Theta^S(t)-\alpha_S\ln S\big]S\frac{\partial U}{\partial S}\\
&-\big[\Theta^Z(t)-\alpha_Z\ln Z\big]Z\frac{\partial U}{\partial Z}-\frac{Z-M}{t-T_l}\frac{\partial U}{\partial M}-\frac{1}{2}\sigma_S^2S^2\frac{\partial^2U}{\partial S^2}\\
&-\frac{1}{2}\sigma_Z^2Z^2\frac{\partial^2U}{\partial Z^2}-\rho\sigma_S\sigma_ZSZ\frac{\partial^2U}{\partial S\partial Z}\\
&-\max_{q\in[0,\bar{q}]}\left\{q\cdot(S-I)+q\cdot\frac{\partial U}{\partial Q}\right\}+\mathcal{O}(\Delta t)\\
=&-\mathcal{L}U+\mathcal{O}(\Delta t). \qquad (4.78)
\end{aligned}$$

Passing the limit of the above equation (4.78), we have the desired result (4.56).

For other cases in Table 4.6, the proof can be done in a similar way. Now, we can conclude with the following theorem:

Theorem 4.1. *Under Assumption 4.1, if the value of $\mathcal{M}_{n,m,c}$ in* (4.55) *is one of the values in the running average vector $\mathbf{M}_{n+1}$, that is,*

$$\mathfrak{d}_{n+1}(\hat{m})=\mathcal{M}_{n,m,c}$$

for some non-negative integer $\hat{m} < H$, the tree method (4.55) *is consistent with the HJB equation* (4.50) *in the order of Δt.*

Now we take into account the linear interpolation (4.32). Since it is hard to measure the interpolation error when using the non-uniform grid, we take the further discretization as an example. In Case 1, denote the interpolated value of U by $\tilde{U}$. Then (4.58) is given by

$$\begin{aligned}\mathcal{U}(t,S,Z,M,I,Q) =& P_{uu}\cdot\tilde{U}\big(t,S^u,Z^u,M^u,I,Q+q\cdot\Delta t\big)+\\ & P_{um}\cdot\tilde{U}\big(t,S^u,Z^m,M^m,I,Q+q\cdot\Delta t\big)+\\ & P_{ud}\cdot\tilde{U}\big(t,S^u,Z^d,M^d,I,Q+q\cdot\Delta t\big)+\\ & P_{mu}\cdot\tilde{U}\big(t,S^m,Z^u,M^u,I,Q+q\cdot\Delta t\big)+\\ & P_{mm}\cdot\tilde{U}\big(t,S^m,Z^m,M^m,I,Q+q\cdot\Delta t\big)+\\ & P_{md}\cdot\tilde{U}\big(t,S^m,Z^d,M^d,I,Q+q\cdot\Delta t\big)+\\ & P_{du}\cdot\tilde{U}\big(t,S^d,Z^u,M^u,I,Q+q\cdot\Delta t\big)+\\ & P_{dm}\cdot\tilde{U}\big(t,S^d,Z^m,M^m,I,Q+q\cdot\Delta t\big)+\\ & P_{dd}\cdot\tilde{U}\big(t,S^d,Z^d,M^d,I,Q+q\cdot\Delta t\big),\end{aligned}\tag{4.79}$$

Recall the further discretization (4.19). For linear interpolation, we have

$$\tilde{U}-U=\mathcal{O}(\Delta M^2).\tag{4.80}$$

Together with (4.75), we have

$$\begin{aligned}&e^{-r\Delta t}\mathcal{U}(t,S,Z,M,I,Q)\\ =&U\big(t,S,Z,M,I,Q\big)+\Big\{\big[\Theta^S(t)-\alpha_S\ln S\big]S\frac{\partial U}{\partial S}+\big[\Theta^Z(t)-\alpha_Z\ln Z\big]Z\frac{\partial U}{\partial Z}\\ &+\frac{Z-M}{t-T_l}\frac{\partial U}{\partial M}+q\cdot\frac{\partial U}{\partial Q}+\frac{1}{2}\sigma_S^2S^2\frac{\partial^2U}{\partial S^2}+\frac{1}{2}\sigma_Z^2Z^2\frac{\partial^2U}{\partial Z^2}\\ &+\rho\sigma_S\sigma_ZSZ\frac{\partial^2U}{\partial S\partial Z}-rU\big(t,S,Z,M,I,Q\big)\Big\}\Delta t+\mathcal{O}(\Delta t^2+\Delta M^2)\end{aligned}$$

It follows that

$$\begin{aligned}&\frac{1}{\Delta t}(U - F_{\Delta t}U)(t-\Delta t, S, Z, M, I, Q)\\ =&-\mathcal{L}U + \mathcal{O}(\Delta t + \frac{\Delta M^2}{\Delta t}).\end{aligned} \tag{4.81}$$

Recall that

$$\Delta M = \frac{\Delta Z}{F} = \frac{\sigma_Z\sqrt{3\Delta t}}{F},$$

if $F = \mathcal{O}(\Delta t^{-\frac{1}{2}})$,

$$\mathcal{O}(\Delta t + \frac{\Delta M^2}{\Delta t}) = \mathcal{O}(\Delta t).$$

That is, if $F = \mathcal{O}(\Delta t^{-\frac{1}{2}})$, passing the limit of equation (4.81), we can have the desired result (4.56).

For other cases in Table 4.6, the proof can be done in a similar way. Now, we can conclude with the following theorem:

Theorem 4.2. *Under Assumption 4.1, when using further discretization (4.19) with* $F = \mathcal{O}(\Delta t^{-\frac{1}{2}})$*, the tree method* (4.55) *is consistent with the HJB equation* (4.50) *in the order of* Δt*.*

Chapter 5

The least-squares Monte Carlo approach

5.1 Introduction

The least-squares Monte Carlo simulation (LSMC, see Longstaff and Schwartz (2001)) is a powerful tool when pricing options with early exercise features. It has been extended to accommodate options with multiple exercise opportunities in Meinshausen and Hambly (2004), Aleksandrov and Hambly (2010) and Bender (2011). More specifically, the LSMC method has been used in Dörr (2003), Barrera-Esteve et al. (2006), Thanawalla (2006), Holden, Løland and Lindqvist (2011) and Bernhart (2011) to price GSA contracts.

In the existing literature, the strike price, or the index, is commonly assumed to be deterministic. Then, it is enough to perform the regression simply on the underlying price. When it comes to the evaluation of GSAs with indexation, in each month, the value of the index is determined by the weighted average price of the crude oil in the previous month. Since the index is surely not deterministic, a common approach is to perform the regression on both the gas price and the index. A simplified approach is also suggested in Holden, Løland and Lindqvist (2011) and Edoli (2013) where the regression is performed on the instant payoff. Since the index is computed by using the crude oil price, however, this should also affect the value of a GSA. This problem is noticed in Bernhart (2011) and Grau (2008), where the authors suggest the regression should also be performed on the past underlying prices which are used to compute the current strike price or index. Two different approaches have been used in these two references. The regression

method in Grau (2008) regresses on all past underlying prices which are used to compute the strike price. This approach leads to a very high-dimensional problem, especially in the evaluation of the GSA contract, since 30 (we assume one month contains 30 days) past oil prices are used to compute the index. Although the author proposes a technique based on a sparse grid to reduce the number of basis functions used, it is still very time consuming and can hardly be feasible when the dimension exceeds ten. Bernhart (2011) uses several so-called Laguerre processes in Bernhart, Tankov and Warin (2011) to approximate the moving average process before performing the regression on these Laguerre processes together with the gas price and the index. This action still leads to a very high-dimensional problem, however, since the approximation the author uses is accurate when many Laguerre processes are used. In addition, as reported in Warin (2012), this approach does not give significant improvement compared with the normal regression on the gas price and the index.

This chapter includes three aspects. Firstly, we build the LSMC algorithm for the purpose of evaluating the GSAs with indexation. Secondly, we investigate the performance and find ways to get better results by using the LSMC algorithm. This is done by comparing the outcomes of the LSMC algorithm by performing the regression on different variables. Thirdly, we provide algorithms which can be used to compare the contract value with the tree algorithm built in Chapter 4.

Since we have used the model

$$\begin{cases} dX(t) = -\alpha_S X(t)dt + \sigma_S dB^S(t) \\ S(t) = e^{X(t)+a(t)} \\ dY(t) = -\alpha_Z Y(t)dt + \sigma_Z dB^Z(t) \\ Z(t) = e^{Y(t)+b(t)} \end{cases}, \tag{5.1}$$

in the tree algorithm (see (4.38)) in Chapter 4, to provide a fair comparison between the tree algorithm and the LSMC algorithm, we continue to use this model in this chapter. Due to the flexible nature of the Monte Carlo simulation, however, the LSMC algorithm can surely accommodate other models. We give a description on how to generate paths by using (5.1) in Remark 5.3.

This chapter is organized as follows: Section 5.2 builds the LSMC algorithm for the evaluation of GSAs with indexation. Section 5.3 refines the LSMC algorithm by introducing the exercising rule. Section 5.4 gives an algorithm which can be used to get the upper bound of the GSA contract value. Section 5.5 provides an algorithm which gives a benchmark contract value when the penalty of the GSA is not applied. Section 5.6 lists the candidate variables which can be used in the least-squares regression. Section 5.7 provides the basis functions we use in the LSMC algorithm. Section 5.8 analyzes the performance of the LSMC algorithm through numerous examples. We draw conclusions in Section 5.9.

5.2 The least-squares Monte Carlo Approach

In this section, we build the least-squares algorithm for the evaluation of the GSA. To ease the notations in this chapter, we denote the gas price, the oil price, the value of the running average and the index at time t_n by S_n, Z_n, M_n and I_n, respectively. In the same sprit as (4.7) and (4.8), we have, for $n = lD + d$, $l = 0, \ldots, L-1$ and $d = 1, \ldots, D$,

$$M_{lD+d} = \frac{1}{d}\sum_{j=1}^{d} Z_{lD+j} \text{ for } l = 0, 1, \ldots, L-2.$$

$$I_{lD+d} = \begin{cases} K & \text{for } l = 0, \\ \frac{1}{D}\sum_{j=1}^{D} Z_{(l-1)D+j} = M_{lD}, & \text{for } l = 1, 2, \ldots, L-1. \end{cases}$$

Recall that L corresponds to the number of months in a year and D corresponds to the number of days in a month. Denote the contract value of the GSA by $V(t_n, S, I, Q)$ where the gas price is S, the index is I and the period to date is Q at time t_n, $n = 1, 2, \ldots, N$. Recall (4.11), we have

$$V(t_n, S, I, Q) = \sup_{q_{t_k} \in [q_{\min}, q_{\max}]} \mathbb{E}\Bigg[\sum_{k=n}^{N} e^{-r(t_k - t_n)} q_{t_k} \Big(S_k - I_k \Big) + e^{-r(T-t_n)} \mathcal{P}\left(I_N, Q_T\right) \Bigg| S_n = S, I_n = I, Q_{t_n} = Q \Bigg]. \quad (5.2)$$

where

$$Q_T = Q + \sum_{k=n}^{N} q_{t_k}.$$

Recall that $Q_{t_n} = \sum_{k=1}^{n-1} q_{t_k}$ for $n \geq 2$ and $Q_{t_1} = 0$. For $t_N = T$, we find the optimal decision $q^* = q^*(t_N, S, I, Q)$ by using (4.26), and the terminal contract value follows at once,

$$V(T, S, I, Q) = q^* \cdot (S - I) + \mathcal{P}(I, Q + q^*). \tag{5.3}$$

From the contract value (5.2), we have

$$\begin{aligned} V(t_n, S, I, Q) = \sup_{q_{t_n}, \ldots, q_{t_N}} \mathbb{E}\Bigg[& q_{t_n}\Big(S - I\Big) + e^{-r\Delta t}\Bigg[\sum_{k=n+1}^{N} e^{-r(t_k - t_{n+1})} q_{t_k}\Big(S_k \\ & - I_k\Big) + e^{-r(T - t_{n+1})}\mathcal{P}\left(I_N, Q_T\right)\Bigg]\Bigg| S_n = S, I_n = I, Q_{t_n} = Q\Bigg] \\ = \sup_{q_{t_n} \in [q_{\min}, q_{\max}]} \mathbb{E}\Bigg[& q_{t_n}\Big(S - I\Big) + e^{-r\Delta t}\cdot \\ & \mathbb{E}\Big[V(t_{n+1}, S_{n+1}, I_{n+1}, Q_{t_n} + q_{t_n}) | S_n = S, I_n = I, Q_{t_n} = Q\Big]\Bigg]. \end{aligned} \tag{5.4}$$

In (5.4), the conditional expectation together with the discount factor

$$e^{-r\Delta t} \cdot \mathbb{E}\Big[V(t_{n+1}, S_{n+1}, I_{n+1}, Q_{t_n} + q_{t_n}) | S_n = S, I_n = I, Q_{t_n} = Q\Big]$$

is called the continuation value. Let $C(t_n, S, I, Q, q)$ be the continuation value where the gas price is S, the index is I, the period to date is Q and the decision is q at time t_n. That is,

$$C(t_n, S, I, Q, q) = e^{-r\Delta t}\mathbb{E}\Big[V(t_{n+1}, S_{n+1}, I_{n+1}, Q + q)\Big| S_n = S, I_n = I\Big]. \tag{5.5}$$

The main idea of the LSMC is that the continuation value can be approximated by the linear combination of a set of basis functions of the current state (the gas price, the oil price, the running average, etc.). These basis functions can be monomials, Laguerre polynomials, splines, radial basis functions, etc..

In our case, at each time t_n, $n = 1, 2, \ldots, N-1$, (5.5) can be approximated through

$$C(t_n, S, I, Q, q) \approx \sum_{k=1}^{K_b} \beta_{n,k}^{Q,q} \phi_{n,k}(\mathcal{E}_n) = \hat{C}_{K_b}(t_n, S, I, Q, q),$$

where K_b is a positive integer. $\hat{C}_{K_b}(t_n, S, I, Q, q)$ is the approximated value of $C(t_n, S, I, Q, q)$. $\beta_{n,k}^{Q,q}$, $k = 1, \ldots, K_b$, is the coefficient associated with the basis function $\phi_{n,k}(\mathcal{E}_n)$ where the period to date equals Q and the daily decision is q at time t_n. We call this the regression coefficient. $\mathcal{E}_n$ is a vector containing some appropriate explanatory variables at time t_n. In the rest of this chapter, we call $\mathcal{E}_n$ the explanatory vector. The explanatory variables are the variables which have impacts on the continuation value. Intuitively speaking, the possible candidate explanatory variables are the gas price, the oil price, the running average and the index. We discuss the choices of explanatory variables in Section 5.6 and explore different possibilities for $\mathcal{E}_n$ in Section 5.8.

Remark 5.1. For the LSMC, the approximated continuation value converges to the continuation value when $K_b \to +\infty$ (see Clément, Lamberton and Protter (2002)). That is

$$C(t_n, S, I, Q, q) = \lim_{K_b \to +\infty} \sum_{k=1}^{K_b} \beta_{n,k}^{Q,q} \phi_{n,k}(\mathcal{E}_n),$$

In practical, however, we can only work with a finite K_b. Fortunately, for options with early exercises features, a not very large number of the basis functions usually returns an acceptable approximation of $C(t_n, S, I, Q, q)$ (see Moreno and Navas (2003), Meinshausen and Hambly (2004) and Stentoft (2004)). More details of the number of basis functions are given in Section 5.8.

In LSMC, the cross-paths information of many simulated paths is used to approximate the continuation value. Suppose we have G simulated independent paths of the two-dimensional Markov process (S, Z):

$$\left((S_0^{(g)}, Z_0^{(g)}), (S_1^{(g)}, Z_1^{(g)}), (S_2^{(g)}, Z_2^{(g)}), \ldots, (S_N^{(g)}, Z_N^{(g)}) \right),$$

where $g = 1, 2, \ldots, G$. For each path g, we calculate the explanatory vector $\mathcal{E}^{(g)}$ and get G paths of $\mathcal{E}$:

$$\left(\mathcal{E}_0^{(g)}, \mathcal{E}_1^{(g)}, \mathcal{E}_2^{(g)}, \ldots, \mathcal{E}_N^{(g)}\right),$$

where $g = 1, 2, \ldots, G$. Now, denote $\hat{C}_{K_b}^{(g)}(t_n, S^{(g)}, I^{(g)}, Q, q)$ as the approximated continuation value at time t_n of the gth realization (path) obtained by the linear combination of basis functions $\phi_k(\mathcal{E}_n^{(g)})$, $k = 1, \ldots, K_b$. That is,

$$\hat{C}_{K_b}^{(g)}(t_n, S^{(g)}, I^{(g)}, Q, q) = \sum_{k=1}^{K_b} \beta_{n,k}^{Q,q} \phi_k(\mathcal{E}_n^{(g)})$$

We use the same set of basis functions at all times t_n, $n = 1, \ldots, N$, which means we can drop the subscription n in $\phi_{n,k}$. Similarly, let $\hat{V}^{(g)}(t_n, S^{(g)}, I^{(g)}, Q)$ be the approximated contract value at time t_n of the gth realization calculated through $\hat{C}_{K_b}^{(g)}(t_n, S^{(g)}, I^{(g)}, Q, q)$, which is given by

$$\hat{V}^{(g)}(t_n, S^{(g)}, I^{(g)}, Q) = \sup_{q \in [q_{\min}, q_{\max}]} \left[q \cdot \left(S^{(g)} - I^{(g)} \right) + \hat{C}_{K_b}^{(g)}(t_n, S^{(g)}, I^{(g)}, Q, q) \right].$$

Given K_b basis functions, let $\beta_{n,k}^{Q,q,*}$ be the optimal regression coefficient associated with $\phi_k(\mathcal{E}_n)$. To obtain $\beta_{n,k}^{Q,q,*}$, we perform a least-squares regression by solving

$$\begin{aligned} &\min_{\beta_{n,1}^{Q,q}, \ldots, \beta_{n,K_b}^{Q,q}} \sum_{g=1}^{G} \left[e^{-r\Delta t} \hat{V}^{(g)}(t_{n+1}, S^{(g)}, I^{(g)}, Q+q) - \hat{C}_{K_b}^{(g)}(t_n, S^{(g)}, I^{(g)}, Q, q) \right]^2 \\ = &\min_{\beta_{n,1}^{Q,q}, \ldots, \beta_{n,K_b}^{Q,q}} \sum_{g=1}^{G} \left[e^{-r\Delta t} \hat{V}^{(g)}(t_{n+1}, S^{(g)}, I^{(g)}, Q+q) - \sum_{k=1}^{K_b} \beta_{n,k}^{Q,q} \phi_k(\mathcal{E}_n^{(g)}) \right]^2 \end{aligned} \tag{5.6}$$

Note that, the optimal $\beta_{n,k}^{Q,q,*}$ depends on both the period to date Q and the daily decision q. That is, at each time t_n, $n = 1, \ldots, N-1$, we find the optimal coefficients for all possible combinations of Q and q. Once we have $\beta_{n,k}^{Q,q,*}$, we compute the continuation value of each path g using these optimal coefficients. Then for each path g, for each fixed period to date Q, the contract

value of path g at time t_n is given by

$$\hat{V}^{(g)}(t_n, S^{(g)}, I^{(g)}, Q) = \sup_{q\in[q_{\min},q_{\max}]} \left[q\Big(S^{(g)} - I^{(g)}\Big) + \sum_{k=1}^{K_b} \beta_{n,k}^{Q,q,*} \phi_k(\mathcal{E}_n^{(g)}) \right]. \quad (5.7)$$

In addition, at time t_N, by using (5.3), we can compute the contract value of each path g through

$$\hat{V}(g)(t_N, S^{(g)}, I^{(g)}, Q) = q^* \cdot (S^{(g)} - I^{(g)}) + \mathcal{P}(I^{(g)}, Q + q^*), \quad (5.8)$$

where the optimal decision $q^* = q^*(t_N, S^{(g)}, I^{(g)}, Q)$ is found by using (4.26).

The LSMC can be interpreted as follows: First, we simulate G independent sample paths of the underlying processes (S, Z) which gives the explanatory vector of each path, and chose a set of K_b basis functions ϕ_k, $k = 1, \ldots, K_b$. Then, at maturity $t_N = T$, the contract value of each path g, $g = 1, \ldots, G$, is calculated through (5.8). Then we work backwards in time. At time t_n, $n = N-1, N-2, \ldots, 1$, for each path g, for all possible periods to date Q, we need to decide how much gas the buyer should purchase to maximize the sum of the instant payoff and the continuation value. This is done by using the information of all simulated paths at time t_{n+1} to find the optimal coefficients $\beta_{n,k}^{Q,q,*}$ (see (5.6)). Then the desired contract value for each path is obtained through (5.7). This procedure is repeated backwards in time until time t_1. Since there is no exercise opportunity at time t_0, the contract value then reads

$$V_0 \approx e^{-r\Delta t} \frac{1}{G} \sum_{g=1}^{G} \hat{V}^{(g)}(t_1, S^{(g)}, I^{(g)}, 0).$$

Remark 5.2. The regression coefficients $\beta_{n,k}^{Q,q,*}$ provide an exercising rule. In (5.7), the contract value is obtained by finding the decision q which maximizes the sum of the instant payoff and the approximated continuation value. Then, at each time t_n, for a given period to date Q and any realization of the underlying (S, Z), we can find the optimal decision $q^* = q^*(t_n, S, I, Q)$

through

$$q^*(t_n, S, I, Q) = \text{argmax}_{q\in[q_{\min},q_{\max}]}\left[q\Big(S-I\Big) + \sum_{k=1}^{K_b}\beta_{n,k}^{Q,q,*}\phi_k(\mathcal{E}_n)\right]. \quad (5.9)$$

As we can see in (5.9), once we have the period to date Q and the realization of the underlying (S, Z), the exercising rule solely depends on the optimal coefficients $\beta_{n,k}^{Q,q,*}$, $k = 1, \ldots, K_b$, since the continuation value is calculated through these coefficients.

Remark 5.3. In this thesis, the underlying processes S and Z are simulated using the Euler scheme (see Maruyama (1955)). With $X_0 = Y_0 = 0$,

$$\begin{aligned} X_n =& X_{n-1} - \alpha_S X_{n-1}\Delta t + \sigma_S \Delta B_n^S, \\ Y_n =& Y_{n-1} - \alpha_Z Y_{n-1}\Delta t + \sigma_Z \Delta B_n^Z, \end{aligned}$$

$n = 1, 2, \ldots, N$, where $\Delta B_n^S = B^S(t_n) - B^S(t_{n-1})$ and $\Delta B_n^Z = B^Z(t_n) - B^Z(t_{n-1})$ are the Brownian increments. Then the realizations of S and Z at each time t_n, $n = 0, 1, \ldots, N$, are given by

$$S_n = e^{X_n + a_n}, \quad Z_n = e^{Y_n + b_n},$$

where a_n and b_n are computed using the tree building procedures. Once we have the realizations of S and Z, we can compute the explanatory variables. As we mentioned before, the possible candidate explanatory variables are the gas price S, the oil price Z, the running average M and the index I. Since we have the realizations of S and Z, M and I can be computed by (4.7) and (4.8), respectively.

We now present the numerical implementation in Algorithm I.

Algorithm I

1. Generate G independent paths of the underlying processes (S, Z) and then compute the explanatory vector $\mathcal{E}$ for each path,

$$\left(\mathcal{E}_0^{(g)}, \mathcal{E}_1^{(g)}, \mathcal{E}_2^{(g)}, \ldots, \mathcal{E}_N^{(g)}\right).$$

where $g = 1, 2, \ldots, G$.

2. At time $t_N = T$, for each path g, for each possible period to date Q, set the terminal contract value by

$$\hat{V}^{(g)}(t_N, S^{(g)}, I^{(g)}, Q) = q^* \cdot (S^{(g)} - I^{(g)}) + \mathcal{P}(I^{(g)}, Q + q^*), \qquad (5.10)$$

where $q^* = q^*(t_N, S^{(g)}, I^{(g)}, Q)$ is obtained by using (4.26).

3. For $n = N - 1, N - 2, \ldots, 1$, by backwards induction, for each possible period to date Q,

   (a) for each possible daily decision q,

      i. find the optimal coefficients $\beta_{n,k}^{Q,q,*}$ by solving (5.6).

      ii. for each path g, calculate the approximated continuation value by

$$\hat{C}_{K_b}^{(g)}(t_n, S^{(g)}, I^{(g)}, Q, q) = \sum_{k=1}^{K_b} \beta_{n,k}^{Q,q,*} \phi_k(\mathcal{E}_n^{(g)}) \qquad (5.11)$$

   (b) for each path g, find the approximated contract value associated with the period to date Q by (5.7).

4. For $n = 0$, for each path g, the contract value $\hat{V}_0^{(g)}$ is

$$\hat{V}_0^{(g)} = e^{-r\Delta t} \hat{V}^{(g)}(t_1, S^{(g)}, I^{(g)}, 0).$$

5. The approximated contract value is

$$\hat{V}_0 = \frac{1}{G} \sum_{g=1}^{G} \hat{V}_0^{(g)}.$$

5.3 LSMC using an exercising rule

When we compute the optimal coefficients (which give us an exercising rule, see Remark 5.2) by solving the minimization problem (5.6), at time t_n, we have assumed the knowledge of $\hat{V}^{(g)}(t_{n+1}, S^{(g)}, I^{(g)}, Q+q)$. At the same time, we are applying this exercising rule to the same set of paths in order to get the

contract values, and this causes a high bias of the contract value (see, for instance, Section 8 in Glasserman (2003) and Broadie and Glasserman (2004)). To tackle this issue, one should simulate a second independent set of independent paths and apply the exercising rule obtained through the first set of paths to this newly generated set of paths through a forward induction. We call this algorithm the LSMC using an exercising rule.

The LSMC using an exercising rule contains two steps. In the first step, given a set of K_b basis functions, at each time t_n, $n = 1, 2, \ldots, N-1$, we find the optimal coefficients $\beta_{n,k}^{Q,q,*}$, $k = 1, \ldots, K_b$, by using Algorithm I for all possible Q and q. In the second step, we generate a second set of paths and apply the exercising rule (see (5.9)) through a forward induction. In the rest of this thesis, the first step and the second step are called the backward scheme and the forward scheme, respectively.

The bang-bang consumption In this chapter, we again assume that our GSAs have the bang-bang consumption (see Theorem 2.2). Under the bang-bang consumption, and letting $q_{\min} = 0$ and $q_{\max} = 1$ (see Remark 2.6), at each time t_n, for each possible period to date Q, we only need to find the optimal coefficients $\beta_{n,k}^{Q,0,*}$ and $\beta_{n,k}^{Q,1,*}$, and hence calculate the approximated continuation value $\hat{C}_{K_b}^{(g)}(t_n, S^{(g)}, I^{(g)}, Q, 0)$ and $\hat{C}_{K_b}^{(g)}(t_n, S^{(g)}, I^{(g)}, Q, 1)$. Furthermore, based on the same reasons, at time t_n, the possible periods to date Q are integers from 0 to $n-1$ (see Remark 2.6). That is, (5.7) becomes

$$\hat{V}^{(g)}(t_n, S^{(g)}, I^{(g)}, Q) = \max \left\{ \sum_{k=1}^{K_b} \beta_{n,k}^{Q,0,*} \phi_k(\mathcal{E}_n^{(g)}), \right.$$
$$\left. \left(S^{(g)} - I^{(g)}\right) + \sum_{k=1}^{K_b} \beta_{n,k}^{Q,1,*} \phi_k(\mathcal{E}_n^{(g)}) \right\}. \tag{5.12}$$

Now, we present the LSMC with an exercising rule in Algorithm II.

Algorithm II

1. Generate G_1 independent paths of the underlying processes (S, Z) and then compute the explanatory vector $\mathcal{E}$ for each path,

$$\left(\mathcal{E}_0^{(g)}, \mathcal{E}_1^{(g)}, \mathcal{E}_2^{(g)}, \ldots, \mathcal{E}_N^{(g)}\right).$$

   where $g = 1, 2, \ldots, G_1$.

2. At time $t_N = T$, for $g = 1, 2, \ldots, G_1$ and $Q = 0, 1, 2, \ldots, N-1$, set the terminal contract value $V^{(g)}(t_N, S^{(g)}, I^{(g)}, Q)$ by (5.10).

3. At time t_n, $n = N-1, N-2, \ldots, 1$. By backwards induction, at each t_n, for $Q = 0, 1, 2, \ldots, n-1$,

   (a) for $q = 0$ and $q = 1$,

      i. find the optimal coefficients $\beta_{n,k}^{Q,q,*}$ by solving (5.6) (replace G with G_1 in (5.6)).

      ii. for $g = 1, \ldots, G_1$, calculate the approximated continuation value by (5.11). That is, calculate $\hat{C}_{K_b}^{(g)}(t_n, S^{(g)}, I^{(g)}, Q, q)$.

   (b) for $g = 1, \ldots, G_1$, find the approximated contract value associated with the period to date Q by

$$\hat{V}^{(g)}(t_n, S^{(g)}, I^{(g)}, Q) = \max\Bigg\{\hat{C}_{K_b}^{(g)}(t_n, S^{(g)}, I^{(g)}, Q, 0), \\ \left(S^{(g)} - I^{(g)}\right) + \hat{C}_{K_b}^{(g)}(t_n, S^{(g)}, I^{(g)}, Q, 1)\Bigg\}. \quad (5.13)$$

4. Generate a second set of G_2 independent paths of the underlying processes (S, Z) and then compute the explanatory vector $\mathcal{E}$ for each path,

$$\left(\mathcal{E}_0^{(g)}, \mathcal{E}_1^{(g)}, \mathcal{E}_2^{(g)}, \ldots, \mathcal{E}_N^{(g)}\right).$$

   where $g = 1, 2, \ldots, G_2$.

5. For each path $g = 1, 2, \ldots, G_2$, perform a forward induction,

   (a) at time t_0, set $Q_1^{(g)} = 0$ and proceed to time t_1.

(b) at time t_n, $n = 1, 2, \ldots, N-1$,

i. if we have

$$\hat{C}^{(g)}_{K_b}(t_n, S^{(g)}, I^{(g)}, Q_{t_n}, 0) < \Big(S^{(g)} - I^{(g)}\Big) + \hat{C}^{(g)}_{K_b}(t_n, S^{(g)}, I^{(g)}, Q_{t_n}, 1),$$

where

$$\hat{C}^{(g)}_{K_b}(t_n, S^{(g)}, I^{(g)}, Q_{t_n}, 0) = \sum_{k=1}^{K_b} \beta_{n,k}^{Q_{t_n},0,*} \phi_k(\mathcal{E}_n^{(g)})$$

and

$$\hat{C}^{(g)}_{K_b}(t_n, S^{(g)}, I^{(g)}, Q_{t_n}, 1) = \sum_{k=1}^{K_b} \beta_{n,k}^{Q_{t_n},1,*} \phi_k(\mathcal{E}_n^{(g)}),$$

set $q^{(g)}_{t_n} = 1$.

ii. otherwise, set $q^{(g)}_{t_n} = 0$.

iii. set $Q^{(g)}_{t_{n+1}} = Q^{(g)}_{t_n} + q^{(g)}_{t_n}$.

(c) at time t_N,

i. find $q^{(g)}_{t_N} = q(t_N, S^{(g)}, I^{(g)}, Q^{(g)}_{t_N})$ by (4.26).

ii. get the penalty $\mathcal{P}^{(g)}(I^{(g)}_N, Q^{(g)}_{t_N})$ by (4.6).

(d) the contract value of path g is given by

$$\hat{V}_0^{(g)} = \sum_{n=1}^{N} e^{-r\cdot n\cdot \Delta t} \cdot q^{(g)}_{t_n} \Big(S^{(g)}_n - I^{(g)}_n\Big) + e^{-rT} \cdot \mathcal{P}^{(g)}(I^{(g)}_N, Q^{(g)}_{t_N}).$$

6. The contract value is

$$\hat{V}_0 = \frac{1}{G_2} \sum_{g=1}^{G_2} \hat{V}_0^{(g)}$$

and the standard error is given by

$$SE(\hat{V}_0) = \frac{\sqrt{\frac{1}{G_2-1}\sum_{g=1}^{G_2}(\hat{V}_0^{(g)} - \hat{V}_0)^2}}{G_2}.$$

Remark 5.4. Recall that, under the bang-bang consumption, we find $\beta_{n,k}^{Q,0,*}$ and $\beta_{n,k}^{Q,1,*}$ for all possible periods to date Q at time t_n via (5.6). If the current period to date is Q, we find $\beta_{n,k}^{Q,0,*}$ by

$$\min_{\beta_{n,1}^{Q,0},\ldots,\beta_{n,k}^{Q,0}} \sum_{g=1}^{G} \left[e^{-r\Delta t} \hat{V}^{(g)}(t_{n+1}, S^{(g)}, I^{(g)}, Q+0) - \sum_{k=1}^{K_b} \beta_{n,k}^{Q,0} \phi_k(\mathcal{E}_n^{(g)}) \right]^2. \quad (5.14)$$

If the current period to date is $Q-1$, we find $\beta_{n,k}^{Q-1,1,*}$ by

$$\min_{\beta_{n,1}^{Q-1,1},\ldots,\beta_{n,k}^{Q-1,1}} \sum_{g=1}^{G} \left[e^{-r\Delta t} \hat{V}^{(g)}(t_{n+1}, S^{(g)}, I^{(g)}, Q-1+1) - \sum_{k=1}^{K_b} \beta_{n,k}^{Q-1,1} \phi_k(\mathcal{E}_n^{(g)}) \right]^2. \quad (5.15)$$

A simple comparison between (5.14) and (5.15) gives

$$\beta_{n,k}^{Q,0,*} = \beta_{n,k}^{Q-1,1,*}. \quad (5.16)$$

Under (5.16), together with (5.11), the following result follows at once:

$$\hat{C}_{K_b}^{(g)}(t_n, S^{(g)}, I^{(g)}, Q, 0) = \hat{C}_{K_b}^{(g)}(t_n, S^{(g)}, I^{(g)}, Q-1, 1).$$

(5.16) gives us the opportunity to save a lot of computing time since it reduces nearly half of the minimization problem (5.6), although it needs to be processed with extra care when we use the parallel computing technique. Without (5.16), at each time t_n, the minimization problem (5.6) can be solved independently for each possible period to date Q. When we apply (5.16), this is not possible since we cannot update $\beta_{n,k}^{Q,0,*}$ until we find $\beta_{n,k}^{Q-1,1,*}$. We give a simple example of how to apply (5.16) using parallel computing in Example 5.1.

Example 5.1. At time t_n, the possible period to date can be $Q \in \{0, 1, \ldots, n-1\}$. Suppose we want to implement our LSMC algorithm using M threads, thread 0, ..., thread $M-1$. Without using (5.16), we can simply put the minimization problem (5.6) of any possible Q into any of those threads. When using (5.16), however, we have to manually divide $\{0, 1, \ldots, n-1\}$ into M consecutive parts: $\{0, \ldots, \bar{Q}\}$, $\{\bar{Q}+1, \ldots, 2\bar{Q}\}$, ..., $\{(M-1)\bar{Q}+1, \bar{Q}+2, \ldots, n-1\}$ and put them into thread 0, ..., thread $M-1$, respectively. $\bar{Q}$ can be set to be the largest integer smaller than $\frac{n-1}{M}$. In each thread, we find $\beta_{n,k}^{Q,0,*}$ $\beta_{n,k}^{Q,1,*}$

through forward induction. Since we update $\beta_{n,k}^{Q,0,*}$ using $\beta_{n,k}^{Q-1,1,*}$, we have to solve the minimization problem (5.6) for both $\beta_{n,k}^{Q,0,*}$ and $\beta_{n,k}^{Q,1,*}$ when

$$Q \in \{0, \bar{Q}+1, 2\bar{Q}+1, 3\bar{Q}+1, \ldots, (M-1)\bar{Q}+1\}.$$

That is, we do not apply (5.16) for the starting period to date in each thread. The reason we present this example is to show that, when we have many threads available (M is sufficiently large), we may gain less benefit from (5.16). Since to the author's knowledge, however, most modern PCs and laptops contain CPUs which have 4 cores, then this issue is not a serious problem.

Remark 5.5. The minimization problem (5.6) can also be written in the form of

$$\boldsymbol{\phi}_n \boldsymbol{\beta}_n^{Q,q} = \boldsymbol{V}_n^q, \tag{5.17}$$

where $\boldsymbol{\phi}_n$ is a $G_1 \times K_b$ matrix given by

$$\boldsymbol{\phi}_n = \begin{pmatrix} \phi_1(\mathcal{E}_n^{(1)}) & \phi_2(\mathcal{E}_n^{(1)}) & \cdots & \phi_{K_b}(\mathcal{E}_n^{(1)}) \\ \phi_1(\mathcal{E}_n^{(2)}) & \phi_2(\mathcal{E}_n^{(2)}) & \cdots & \phi_{K_b}(\mathcal{E}_n^{(2)}) \\ \cdots & \cdots & \cdots & \cdots \\ \phi_1(\mathcal{E}_n^{(G_1)}) & \phi_2(\mathcal{E}_n^{(G_1)}) & \cdots & \phi_{K_b}(\mathcal{E}_n^{(G_1)}) \end{pmatrix}, \tag{5.18}$$

$\boldsymbol{\beta}_n^{Q,q}$ is a $K_b \times 1$ vector given by

$$\boldsymbol{\beta}_n^{Q,q} = (\beta_{n,1}^{Q,q}, \beta_{n,1}^{Q,q}, \ldots, \beta_{n,K_b}^{Q,q})^{\text{trans}}.$$

and $\boldsymbol{V}_n$ is a $G_1 \times 1$ vector given by

$$\boldsymbol{V}_n^q = \begin{pmatrix} e^{-r\Delta t}\hat{V}^{(1)}(t_{n+1}, S^{(1)}, I^{(1)}, Q+q) \\ e^{-r\Delta t}\hat{V}^{(2)}(t_{n+1}, S^{(2)}, I^{(2)}, Q+q) \\ \cdots \\ e^{-r\Delta t}\hat{V}^{(G_1)}(t_{n+1}, S^{(G_1)}, I^{(G_1)}, Q+q) \end{pmatrix}.$$

That is, we seek the optimal regression coefficients $\beta_n^{Q,q,*}$ by solving the over-determined system (5.17) by least-squares regression. In the rest of this thesis,

we denote $\boldsymbol{\beta}_n^{Q,q,*}$ the optimal coefficient vector which is given by

$$\boldsymbol{\beta}_n^{Q,q,*} = (\beta_{n,1}^{Q,q,*}, \beta_{n,1}^{Q,q,*}, \ldots, \beta_{n,K_b}^{Q,q,*})^{\text{trans}}.$$

In addition, as mentioned in Remark 5.4, although the minimization problem (5.6) of each possible period to date cannot be solved independently after applying (5.16), we can always use the parallel computing technique in the construction of these vectors and matrix.

5.4 An upper bound of the GSA

Edoli (2013) proposes an algorithm for the upper bound of the GSA, which is called the Naive Monte Carlo with Linear Programming (NMCLP). The idea is quite straightforward. This algorithm is similar to the forward scheme. Instead of using an exercising rule obtained by the backward scheme, NMCLP maximizes the contract value of each path by linear programming using a deterministic algorithm. Of course, at the same time, both the global and daily constraints are imposed. Then, this upper bound is obtained via the average of all maximized contract values of all paths. We present NMCLP in Algorithm III.

Algorithm III

1. Generate G independent paths of the underlying processes (S, Z)

$$\left((S_0^{(g)}, Z_0^{(g)}), (S_1^{(g)}, Z_1^{(g)}), (S_2^{(g)}, Z_2^{(g)}), \ldots, (S_N^{(g)}, Z_N^{(g)})\right),$$

   where $g = 1, 2, \ldots, G$. And calculate the index values $I_n^{(g)}$, $n = 1, \ldots, N$, for each path g.

2. Find $V^{(g)}(0)$ for each path g through

$$\begin{aligned}
\underset{\mathbf{q},x_1,x_2}{\text{maximize}} \quad & V^{(g)}(0) = \sum_{n=1}^{N} e^{-rt_n} q_{t_n} (S_n^{(g)} - I_n^{(g)}) - I_N^{(g)} \cdot x_1 - I_N^{(g)} \cdot x_2 \\
\text{subject to} \quad & Q_T = \sum_{n=1}^{N} q_{t_n}, \\
& MB - Q_T \leq x_1 \leq MB, \\
& Q_T - ACQ \leq x_2 \leq N - ACQ, \\
& x_1 \geq 0, \\
& x_2 \geq 0, \\
& q_{t_n} \in \{0, 1\},\ n = 1, 2, \ldots, N.
\end{aligned} \tag{5.19}$$

3. The upper bound reads

$$\hat{V}_0 = \frac{1}{G} \sum_{g=1}^{G} V^{(g)}(0).$$

Actually, the above algorithm is fast and easy to implement using a deterministic algorithm. One suggestion is to use the built-in Matlab function `intlinprog()`. In addition, the above algorithm gives the upper bound of the GSA value in the sense that it has the so-called perfect foresight of the future movements of the underlying prices. That is, at time t_0, the contract value $V^{(g)}(0)$ of each path g is calculated via the maximization with the knowledge of all price realizations in all future time steps t_n, $n = 1, 2, \ldots, N$. Then, the value $V^{(g)}(0)$ is obtained in a deterministic environment instead of an uncertain environment. The perfect foresight can significantly overprice the contract value and hence gives an upper bound.

Next, since Edoli (2013) does not elaborate how the global and daily constraints are imposed in (5.19), we give some explanations here. Obviously, since we have assumed the bang-bang consumption, $q_{t_n} \in \{0, 1\}$ in (5.19) has

imposed the daily constraints $0 \leq q_{t_n} \leq 1$. The part

$$Q_T = \sum_{n=1}^{N} q_{t_n}, \tag{5.20}$$

$$MB - Q_T \leq x_1 \leq MB, \tag{5.21}$$

$$Q_T - ACQ \leq x_2 \leq N - ACQ, \tag{5.22}$$

$$x_1 \geq 0,$$

$$x_2 \geq 0$$

makes sure the global constraints are applied. (5.20) means Q_T is the total volume of gas taken. Recall (2.3) and Remark 2.1, $Q_T \leq N$, so the inequality (5.22) always holds. Here we recall the penalty function

$$\mathcal{P}\left(I_N^{(g)}, Q_T\right) = -I_N^{(g)} \cdot \underbrace{\max\left\{MB - Q_T, 0\right\}}_{\text{Part 1}} - I_N^{(g)} \cdot \underbrace{\max\left\{Q_T - ACQ, 0\right\}}_{\text{Part 2}}. \tag{5.23}$$

What we want to do is to use x_1 and x_2 to mimic Part 1 and Part 2 in (5.23), respectively. We have the following scenarios:

1. $0 \leq Q_T < MB$. In this case, the buyer has to pay penalties because of the insufficient gas taken. By (5.23), the penalty is

   $$\mathcal{P}\left(I_N^{(g)}, Q_T\right) = -I_N^{(g)} \cdot (MB - Q_T).$$

   Since $MB - Q_T > 0$ and $Q_T - ACQ < 0$, we have $MB - Q_T \leq x_1 \leq MB$ and $0 \leq x_2 \leq N - ACQ$. Note that, in (5.19), $V^{(g)}(0)$ decreases linearly in both x_1 and x_2. Then, through the linear programming, $x_1 = MB - Q_T$ and $x_2 = 0$, which gives

   $$-I_N^{(g)} \cdot x_1 - I_N^{(g)} \cdot x_2 = -I_N^{(g)} \cdot (MB - Q_T).$$

2. $MB \leq Q_T \leq ACQ$. In this case, no penalty needs to be paid. Since $MB - Q_T < 0$ and $Q_T - ACQ < 0$, we have $0 \leq x_1 \leq MB$ and $0 \leq x_2 \leq N - ACQ$. Through the linear programming, $x_1 = 0$ and $x_2 = 0$, which gives

   $$-I_N^{(g)} \cdot x_1 - I_N^{(g)} \cdot x_2 = 0.$$

3. $ACQ < Q_T \leq N$. In this case, the buyer has to pay penalties because of the excess gas taken. By (5.23), the penalty is

$$\mathcal{P}\left(I_N^{(g)}, Q_T\right) = -I_N^{(g)} \cdot (Q_T - ACQ).$$

Since $MB - Q_T < 0$ and $Q_T - ACQ > 0$, we have $0 \leq x_1 \leq MB$ and $Q_T - ACQ \leq x_2 \leq N - ACQ$. Then, through the linear programming, $x_1 = 0$ and $x_2 = Q_T - ACQ$, which gives

$$-I_N^{(g)} \cdot x_1 - I_N^{(g)} \cdot x_2 = -I_N^{(g)} \cdot (Q_T - ACQ).$$

Another thing worth mentioning is that, since $0 \leq Q_T \leq N$, the maximum penalty related to insufficient gas purchased (when $Q_T = 0$) the buyer would pay is $-I_N^{(g)} \cdot MB$, and the maximum penalty related to excess gas purchased (when $Q_T = N$) the buyer would pay is $-I_N^{(g)} \cdot (N - ACQ)$. This is the reason why we have $x_1 \leq MB$ and $x_2 \leq N - ACQ$ in (5.21) and (5.22), respectively.

5.5 A benchmark when penalties are not involved

It is well-known that there is no explicit formula for the values of options with early exercise features. This is why we seek numerical solutions for our GSA contracts. In this section, we provide an algorithm which gives a trustworthy benchmark of GSAs when penalties are not applied.

Suppose the non-trivial condition does not hold. That is,

$$MB = 0 \text{ and } N \leq ACQ. \tag{5.24}$$

The buyer can freely take gas at any time t_n, $n = 1, 2, \ldots, N$, without worrying about the possible penalties as long as the daily constraints are satisfied. Then, when (5.24) holds, the GSA is equivalent to a strip of N European options covering the same exercisable dates. The n-th European option has the terminal date t_n and the strike price I_n. The payoff of the n-th European option is given by

$$\max\{S_n - I_n, 0\}.$$

Due to the unclear distribution of the index I, however, we still need to find the value of those European options via numerical approaches. We present the algorithm as follows:

Algorithm IV

1. Generate G independent paths of the underlying processes (S, Z) and get
$$\left((S_0^{(g)}, I_0^{(g)}), (S_1^{(g)}, I_1^{(g)}), (S_2^{(g)}, I_2^{(g)}), \ldots, (S_N^{(g)}, I_N^{(g)})\right),$$
where $g = 1, 2, \ldots, G$.

2. For each path g, compute the contract value of path g by
$$V^{(g)}(0) = \sum_{n=1}^{N} e^{-rt_n} \max\{S_n^{(g)} - I_n^{(g)}, 0\}$$

3. Take the expectation over all path, the final contract value reads
$$\hat{V}_0 = \frac{1}{G} \sum_{g=1}^{G} V^{(g)}(0).$$

Similarly to the upper bound algorithm, this benchmark algorithm also gives full knowledge of all future price realizations. This benchmark algorithm does not suffer from the perfect foresight, however. Without worrying about possible future penalties, the buyer only make an "exercise or not" decision based on the current realization of the underlying prices. That is, in each path g, the decision made at time t_n does not affect the decision which is going to be made at time t_{n+1}. In addition, Algorithm IV (point 3.) takes the expectation over all the paths. If the number of paths G is sufficiently large, based on the strong law of large numbers, $\hat{V}_0$ should converge to the true contract value.

The value obtained through this benchmark algorithm cannot be used as a benchmark for Algorithm II, however. This is because, under (5.24), these

two algorithms are equivalent. Since (5.24) holds, by (5.10),

$$V^{(g)}(t_N, S^{(g)}, I^{(g)}, Q) = \max\{S_n^{(g)} - I_n^{(g)}, 0\}$$

for all possible Q at time t_N, which means the terminal values of all paths are irrelevant to the period to date Q. Then, at time t_{N-1}, when we solve the minimization problem (5.6), all obtained coefficient vectors $\boldsymbol{\beta}_{N+1}^{Q,q,*}$, $q = 0, 1$ and $Q = 0, 1, \ldots, N-2$, contain the same values. It follows that $V^{(g)}(t_{N-1}, S^{(g)}, I^{(g)}, Q)$ for all paths are also irrelevant to the period to date Q. If we work backwards in time, at each time t_n, the coefficient vectors $\boldsymbol{\beta}_n^{Q,q,*}$, $q = 0, 1$ and $Q = 0, 1, \ldots, n-1$, also contain the same values. Then, when we apply the exercising rule $\boldsymbol{\beta}_n^{Q,q,*}$ in the forward scheme, the optimal decisions of each path simply depend on the instant payoffs. The buyer will purchase gas if the instant payoff is positive and not purchase otherwise, which means that, for a fixed set of paths, the forward scheme and the benchmark algorithm return the same result.

5.6 The explanatory variables in the evaluation of GSAs

As we have introduced in the last section, the explanatory variables are the appropriate variables which are found to affect the continuation value. Intuitively speaking, some variables may have strong impacts on the continuation value while some variables may not. Next, we present some choices of explanatory variables and then seek to find the most optimal through various numerical examples.

1. In the view of the value function (5.2), one straightforward option for the explanatory variables is that we use the gas price and the index. That is, $\mathcal{E} = (S, I)$. This is quite understandable since the gas price and the index are the variables in the calculation of everyday payoffs and hence should have a great impact on the continuation value. In this way, (5.6) becomes a two-dimensional fitting problem.

2. Although a two-dimensional fitting is not a considerably challenging computational problem, Holden, Løland and Lindqvist (2011) and Edoli (2013) suggest the explanatory variable $\mathcal{E}$ to be $\mathcal{E} = S - I$ which reduces (5.6) to a one-dimensional problem. Hence, this choice should be the

one with the least computing effort without taking account of robustness and accuracy.

3. Recall (4.5) and (4.8), the index I is related to the running average M in the sense that the value of the index in the current month is the value of the running average at the end of the previous month. The running average can affect the contract value by influencing the index. Hence, we can add one dimension M to $\mathcal{E}$ and have $\mathcal{E} = (S, M, I)$.

4. Recall (4.4) and (4.7). The value of the running average is calculated by using oil prices. So, the oil price should also contribute to the contract value by influencing the running average and hence the index. We can further add one more dimension, Z, to $\mathcal{E}$ and thus have $\mathcal{E} = (S, Z, M, I)$.

5. Since the impact of the running average M on the contract value is originally from the oil price Z, we can let the explanatory vector be $\mathcal{E} = (S, Z, I)$.

In Edoli (2013) and Bernhart (2011), the index has been modelled by using the so-called moving average variable. Denote this moving average variable by A. Based on the notations in Section 4.2 and 4.3, the value of A at time $t \in [0, T]$ is given by

$$A(t) = \frac{1}{\delta} \int_{t-\delta}^{t} Z(u)du \tag{5.25}$$

with the convention $Z(t) = Z(0)$ for $t < 0$. It follows that

$$dA(t) = \frac{1}{\delta}\Big(Z(t) - Z(t-\delta)\Big)dt.$$

That is, A is the moving average of the oil price Z over a time window with the fixed length δ. Then the index I at time $t \in (T_1, T]$ is given by

$$I(t) = A(\varphi(t)), \tag{5.26}$$

where $\varphi(t) = \max\{T_l | T_l < t, l = 1, 2, \ldots, L-1.\}$. Again, in $[T_0, T_1]$, the index is still the constant strike price K. Then we have

$$I(t) = \begin{cases} K & \text{for} \quad t \in [T_0, T_1], \\ A(\varphi(t)) & \text{for} \quad t \in (T_1, T]. \end{cases} \tag{5.27}$$

A simple comparison between (4.4) and (5.26) gives us, for $t \in (T_l, T_{l+1}]$, $l = 1, 2, \ldots, L-1$,

$$A(\varphi(t)) = A(T_l) = M(T_l).$$

It follows that (5.27) is equivalent to (4.3). And of course, we have the value of the moving average at time t_n in discrete time, which is given by

$$A(t_n) = \frac{1}{D} \sum_{j=n-D+1}^{n} Z(t_j).$$

Now, we continue to introduce possible choices of the explanatory variables $\mathcal{E}$.

6. Since we have introduced the moving average variable A, and the index I is connected to A through (5.26), we can let $\mathcal{E} = (S, A, I)$.

7. The moving average A is computed by using the oil prices through (5.25), and thus we can let the oil price be one of the explanatory variables and have $\mathcal{E} = (S, Z, A, I)$.

Remark 5.6. Recall that, in the last month, the value of the index is already realized. So, at this point, the running average M and the moving average A lose their impacts on the continuation value, because both of them contribute to the contract value by influencing the future index. It follows that the oil price Z also loses its impact. That is, in the last month, we do not need to treat the oil price, the running average and the moving average as the explanatory variables. In addition, in the first month, since the index I is a constant, we do not need to include I into our explanatory vector either.

Now, we generalize all the possible choices of explanatory variables we have explained above as follows:

C.1 $\mathcal{E}_n = S_n - I_n$, $n = 1, 2, \ldots, N$.

C.2 $$\mathcal{E}_n = \begin{cases} (S_n, I_n), & n = D+1, \ldots, N \\ S_n, & n = 1, 2, \ldots, D \end{cases}.$$

C.3 $$\mathcal{E}_n = \begin{cases} (S_n, I_n), & n = N-D+1, \ldots, N \\ (S_n, A_n, I_n), & n = D+1, \ldots, N-D \\ (S_n, A_n) & n = 1, 2, \ldots, D \end{cases}.$$

$$\textbf{C.4}\ \ \mathcal{E}_n = \begin{cases} (S_n, I_n), & n = N-D+1, \ldots, N \\ (S_n, Z_n, I_n), & n = D+1, \ldots, N-D \\ (S_n, Z_n) & n = 1, 2, \ldots, D \end{cases}.$$

$$\textbf{C.5}\ \ \mathcal{E}_n = \begin{cases} (S_n, I_n), & n = N-D+1, \ldots, N \\ (S_n, M_n, I_n), & n = D+1, \ldots, N-D \\ (S_n, M_n) & n = 1, 2, \ldots, D \end{cases}.$$

$$\textbf{C.6}\ \ \mathcal{E}_n = \begin{cases} (S_n, I_n), & n = N-D+1, \ldots, N \\ (S_n, Z_n, A_n, I_n), & n = D+1, \ldots, N-D \\ (S_n, Z_n, A_n) & n = 1, 2, \ldots, D \end{cases}.$$

$$\textbf{C.7}\ \ \mathcal{E}_n = \begin{cases} (S_n, I_n), & n = N-D+1, \ldots, N \\ (S_n, Z_n, M_n, I_n), & n = D+1, \ldots, N-D \\ (S_n, Z_n, M_n) & n = 1, 2, \ldots, D \end{cases}.$$

5.7 The choice of basis functions

The main idea of the LSMC algorithm is that the continuation value can be approximated by a linear combination of some basis functions through a least-squares regression. A common choice for these basis functions is the monomials, $\{x^n\}_{n=0}^{\infty}$, $x \in \mathbb{R}$, since they are easy to evaluate. Another common choice is a set of orthogonal polynomials. Table 5.1 gives some common families of orthogonal polynomials. These polynomial functions are univariate functions. Since we have multiple explanatory variables (see Section 5.6), one can form the set of multivariate basis functions by using tensor products of univariate basis functions (see Section 6.12 in Judd (1998)).

Definition 5.1. *Given a basis for functions of the single variable x_i, $\{\phi_n^i(x_i)\}_{n=0}^{\infty}$, $i = 1, 2, \ldots, v$, the tensor product basis Φ of v variables $(x_1, x_2, \ldots, x_v)$ is given by*

$$\Phi = \left\{ \prod_{i=1}^{v} \phi_{n_i}^i(x_i) | n_i = 0, 1, 2, \ldots \right\}.$$

In practice, it is natural to use the finite subset of the full polynomial basis. Suppose for each variable x_i, $i = 1, 2, \ldots, v$, we use the polynomials up to the

TABLE 5.1: Common families of orthogonal polynomials

Family	Weighting function	The interval	Definition
Laguerre	e^{-x}	$[0,\infty)$	$L_n(x)=\frac{e^x}{n!}\frac{d^n x}{x^n}\Big(x^n e^{-x}\Big)$
Legendre	$(1-x^2)^{-\frac{1}{2}}$	$[-1,1]$	$P_n(x)=\frac{(-1)^n}{2^n n!}\frac{d^n}{dx^n}\Big[(1-x^2)^n\Big]$
General Chebyshev	$\Big(1-(2x-1)^2\Big)^{-\frac{1}{2}}$	$[0,1]$	$T_n(x)=\cos\Big(n\cos^{-1}(2x-1)\Big)$
Hermite	e^{-x^2}	$(-\infty,\infty)$	$H_n(x)=(-1)^n e^{x^2}\frac{d^n}{dx^n}\Big(e^{-x^2}\Big)$

Note: The column headed "Weighting function" gives the weighting function which ensures orthogonality. The column headed "The interval" indicates the interval over which the respective family is orthogonal.

degree of h, $h \geq 1$. For instance, if we use Laguerre polynomials as the basis functions, for each variable x_i, we have the basis functions

$$\phi_0^i(x_i) = L_0(x_i), \phi_1^i(x_i) = L_1(x_i), \ldots, \phi_h^i(x_i) = L_h(x_i).$$

Let Υ_h^v be the resulting subset of the tensor product basis Φ, it follows that

$$\Upsilon_h^v = \left\{ \prod_{i=1}^{v} \phi_{n_i}^i(x_i) | n_i = 0, 1, 2, \ldots, h \text{ and } i = 1, \ldots, v \right\}. \tag{5.28}$$

This tensor product basis has a serious drawback: even with a not very high degree h, there will be too many functions in Υ_h^v. As we can see, the total number of functions in (5.28) is $(h+1)^v$. This number $(h+1)^v$ grows rapidly in both the dimension v and the degree h, which makes it very computationally challenging to use Υ_h^v. One way to tackle this issue is to use the complete set of polynomials proposed in Judd (1998).

Definition 5.2. *Given v variables $(x_1, x_2, \ldots, x_v)$, the complete set of polynomials of total degree h in v variables is given by*

$$\mathscr{P}_h^v = \left\{ x_1^{h_1} x_2^{h_2} \cdots x_v^{h_v} | h_n \geq 0, n = 1, 2, \ldots, v \text{ and } \sum_{n=1}^{v} h_n \leq h \right\}. \tag{5.29}$$

TABLE 5.2: Number of functions in $\mathscr{P}_h^v$ and Υ_h^v

Degree h	$\#(\mathscr{P}_h^v)$	$\#(\Upsilon_h^v)$
2	$1+v+\frac{v(v+1)}{2}$	$(2+1)^v$
3	$1+v+\frac{v(v+1)}{2}+v^2+\frac{v(v-1)(v-2)}{6}$	$(3+1)^v$

TABLE 5.3: Parameter values.

$\alpha_S = 3$	$\sigma_S = 0.3$	$\alpha_I = 3$	$\sigma_I = 0.3$	$r = 0.05$
$\rho = 0.5$	$q_{\min} = 0$	$q_{\max} = 1$	$MB = 20$	$ACQ = 30$
$T = 1$	$N = 30$	$D = 6$	$L = 5$	

As reported in Judd (1998), the size of the complete set of polynomials only grows polynomially in dimension v for a fixed degree h. Table 5.2 reports the number of functions in $\mathscr{P}_h^v$ and Υ_h^v for a fixed dimension v. For instance, $\#(\mathscr{P}_2^3) = 10$ while $\#(\Upsilon_2^3) = 27$, $\#(\mathscr{P}_3^3) = 20$ while $\#(\Upsilon_3^3) = 64$. Obviously, for a fixed degree, the complete set of polynomials has the advantage of containing fewer functions and hence requiring less computing effort. In this thesis, we use the complete set of polynomials as the basis functions in our least-squares algorithm.

5.8 Numerical tests on the least-squares Monte Carlo

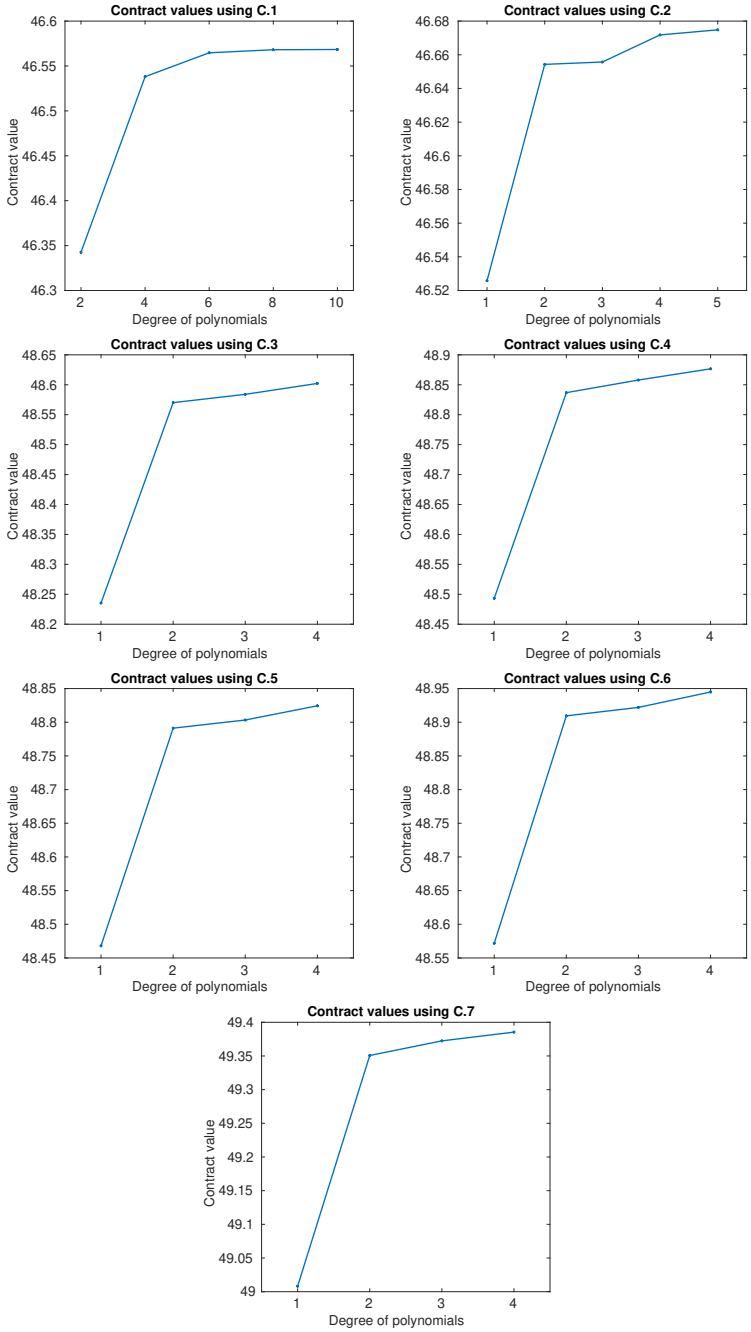
In this section, we investigate the performance of the LSMC through various examples. Unless indicated otherwise, the LSMC algorithm in this section refers to Algorithm II. The CPU we work with is an Intel(R) Core(TM) i7-7700HQ CPU @ 2.80GHz together with 16 GB RAM. All the computing times presented in this section are the times without using the parallel computing technique. In addition, we use (5.16) in all numerical tests in this section. All the numerical examples in this section evaluate a simple GSA with the parameters in Table 5.3.

5.8.1 The impacts of different explanatory vectors

In this subsection, we investigate the appropriate choice of explanatory vectors. The LSMC algorithm (Algorithm II) contains a forward scheme and a backward scheme. The backward scheme provides an exercising rule by using one set of simulated paths, and the forward scheme applies this exercising rule to another set of simulated paths (see Remark 5.2 and Section 5.3). If we fix a set of paths for the backward scheme, given a fixed choice of explanatory vectors, we get different exercising rules by using different degrees of the complete set of polynomials (CSP). If we fix a set of paths for the forward scheme, given a fixed choice of explanatory vectors, we get different contract values by applying different exercising rules. Since the forward scheme is using the information from a fixed set of simulated paths when applying different exercising rules, a better exercising rule would lead to a higher contract value. Before we move to the investigation of the choice of explanatory vectors, we present a numerical example which shows that, for a fixed pair of simulated paths and any given choice of explanatory vectors in Section 5.6, how the degrees of the CSP affect the contract values. Figure 5.1 shows the result. As we can see, for all choices of explanatory vectors, the contract value increases as the degree of the CSP increases. Based on the discussion above, Figure 5.1 indicates that, for a given choice of explanatory vectors, a higher degree of the CSP gives a better exercising rule, and hence a higher contract value. However, for any choice of explanatory vectors, a higher degree of the CSP gives a larger number of basis functions $K_b = \#(\mathscr{P}_h^v)$, where h is the degree of the CSP and v is the number of variables in the explanatory vectors (see (5.29)). Recall the minimization problem (5.6), a larger K_b means more computing effort. Table 5.4 gives the number of basis functions in the CSP for given h and v. Also, to compare the computing effort across different choices of explanatory vectors, one should look into the number of basis functions instead of the degree of the CSP, since a fixed degree of the CSP gives different number of basis functions for different choices of explanatory vectors (see (5.29) and Table 5.4).

Following the discussion above, if we fix a set of simulated paths for the backward scheme, we get different exercising rules by using different

FIGURE 5.1: Contract values w.r.t. the degree of the CSP



Note: For each choice of explanatory vectors in Section 5.6, we simulate a pair of fixed sets of paths, one for the backward scheme and one for the forward scheme. The number of paths used in the backward scheme is $G_1 = 100000$ and the number of paths used in the forward scheme is $G_2 = 1000000$.

TABLE 5.4: Numbers of basis functions in the CSP

$\#(\mathscr{P}_h^v)$	The degree of the CSP h									
	1	2	3	4	5	6	7	8	9	10
$v=1$	2	3	4	5	6	7	8	9	10	11
$v=2$	3	6	10	15	21	28	36	45	55	66
$v=3$	4	10	20	35	56	84	120	165	220	286
$v=4$	5	15	35	70	126	210	330	495	715	1001

Note: $\#(\mathscr{P}_h^v)$ is the number of basis functions in the CSP (see (5.29)).

choices of explanatory vectors. Again, if we fix a set of paths for the forward scheme, we get different contract values by applying different exercising rules obtained from different choices of explanatory vectors. Since the forward scheme is using the same information from a fixed set of paths, a higher contract value would indicate a better exercising rule. That is, with the same computing effort (the same number of basis functions), if a choice of the explanatory vector returns a higher contract value, it indicates that this choice is a better one. However, as we can see in Table 5.4, it is hard to find a common number of basis functions across $v = 1, 2, 3, 4$ (number of variables in the explanatory vector). Fortunately, from Figure 5.1, we know that more computing effort (which means a higher degree of the CSP and more basis functions) gives a better exercising rule for any given choice of the explanatory vector. That is, if we can find a choice of the explanatory vector with less computing effort which still offers a better exercising rule (a higher contract value), it indicates that this choice is a better choice. Following this logic, we design numerical examples in the rest of this subsection. In this subsection, the number of paths used in the backward scheme is $G_1 = 100000$ and the number of paths used in the forward scheme is $G_2 = 1000000$.

To make a fair comparison, we simulate a pair of fixed sets of paths, one for the backward scheme and one for the forward scheme. Then we perform the LSMC on these two sets of paths by using different choices of explanatory vectors. That is, the contract values of using different explanatory vectors are using the same information from the simulated paths in both the backward scheme and forward scheme. In addition, we use the values obtained by

TABLE 5.5: Comparison between **C.1** and **C.2**

EV	Test 1	Test 2	Test 3	Test 4	Test 5
	46.7328	46.5596	46.4432	46.6162	46.5386
C.1	0.0799976	0.0801288	0.0801001	0.079951	0.0800532
	-1.1751%	-1.1901%	-1.2105%	-1.1959%	-1.2035%
	47.2885	47.1204	47.0123	47.1768	47.1055
C.2	0.0799626	0.0800548	0.0800585	0.0799066	0.0800009
	0%	0%	0%	0%	0%

Note: Each cell contains, from top to bottom, the contract value, the standard error and by what proportion this value is lower than the value obtained by using **C.2**.

TABLE 5.6: The number of basis functions used in the tests described in Table 5.7.

	C.2	**C.3**	**C.4**	**C.5**	**C.6**	**C.7**
Degree h	5	3	3	3	2	2
$\#(\mathscr{P}_h^v)$	21	20	20	20	15	15

Note: For **C.2**, the number of variables in the explanatory is $v = 2$. For **C.3**-**C.5**, $v = 3$. For **C.6** and **C.7**, $v = 4$. The degrees h in this table are chosen in the following way: for a fixed pair of simulated paths, using **C.2** in the LSMC algorithm requires the most computing effort. Using **C.3**, **C.4** or **C.5** requires the same computing effort which is less than the effort when using **C.2**. Using **C.6** or **C.7** requires the same computing effort which is the least effort required among using **C.2**-**C.7**

using the explanatory vector **C.2** as benchmark values when we compare different choices of the explanatory vector, since **C.2** is the most commonly used choice. We first eliminate one choice of the explanatory vector by a comparison between the simplest choice **C.1** and our benchmark choice **C.2**. Table 5.5 gives the results. In Table 5.5, for the explanatory vector **C.1**, we use the complete set of polynomials (CSP) of degree 10, which gives us $\#(\mathscr{P}_{10}^1) = 11$ basis functions. For the explanatory vector **C.2**, we use the CSP of degree 2 which gives us $\#(\mathscr{P}_2^2) = 6$ basis functions. As we can see, even with more

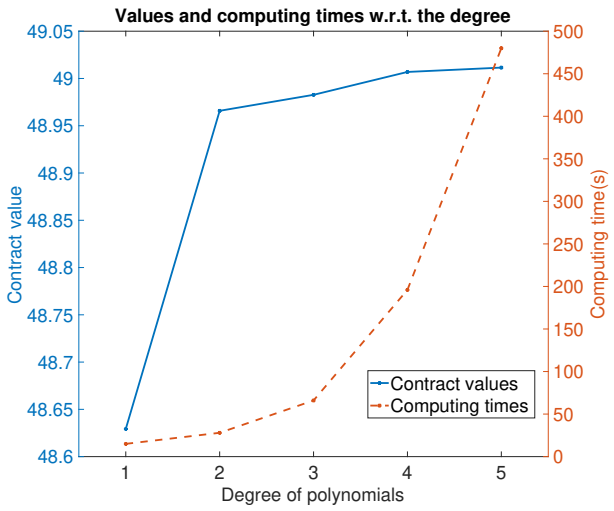
TABLE 5.7: Comparison between explanatory vectors

EV	Test 1	Test 2	Test 3	Test 4	Test 5
	47.0937	46.9721	47.2075	47.1979	47.1045
C.2	0.0800553	0.0799599	0.0800025	0.0799262	0.0799127
	0 %	0%	0%	0%	0%
	48.6803	48.5469	48.7841	48.7775	48.6971
C.3	0.079145	0.0790792	0.0791467	0.0790497	0.0790501
	3.36909%	3.35264%	3.33971%	3.34675%	3.38104%
	48.7515	48.6192	48.8538	48.853	48.7641
C.4	0.0785376	0.0784558	0.0785284	0.0784265	0.0784246
	3.52027%	3.50662%	3.48738%	3.50677%	3.52317%
	48.8363	48.703	48.9461	48.9314	48.8538
C.5	0.0790028	0.0155308	0.0790012	0.0789076	0.078903
	3.70044%	3.68497%	3.68294%	3.67278%	3.71361%
	48.9749	48.8449	49.0763	49.0692	48.9875
C.6	0.0784984	0.0784184	0.0785027	0.0784	0.0783985
	3.99469%	3.98699%	3.95877%	3.96476%	3.99754%
	48.989	48.8594	49.0902	49.0809	49.0006
C.7	0.0784945	0.0784225	0.0785008	0.0784126	0.0783935
	4.0247%	4.01785%	3.98826%	3.98956%	4.02544%

Note: Each cell contains, from top to bottom, the contract value, the standard error and by what proportion this value is higher than the value obtained by using **C.2**.

computing effort ($\#(\mathscr{P}_{10}^1) > \#(\mathscr{P}_2^2)$), the value obtained by using **C.1** is always about 1.2% less than the value obtained by using **C.2**. Table 5.5 indicates that **C.2** is a better choice than **C.1** since it offers a better exercising rule with less computing effort. Now, we perform tests on all the remaining explanatory vectors **C.2**-**C.7**. The numbers of basis functions used together with the degrees of the CSP are given in Table 5.6. The results are shown in Table 5.7. As we can see, even using the largest number of basis functions ($\#(\mathscr{P}_5^2) = 21$), the value obtained by using **C.2** is significantly smaller compared with other choices of explanatory vectors. **C.3**-**C.5** use the same number of basis functions, but **C.5** always returns the largest value and **C.3** always returns the smallest value among these three choices. This evidences that **C.5** is a better choice compared with the other two, and also evidences that the running average M has more impact on the contract value than the moving average A. In addition, even with the smallest number of basis functions ($\#(\mathscr{P}_2^4) = 15$), **C.6** and **C.7** always return larger values compared with

FIGURE 5.2: Values and computing times w.r.t. the degree h



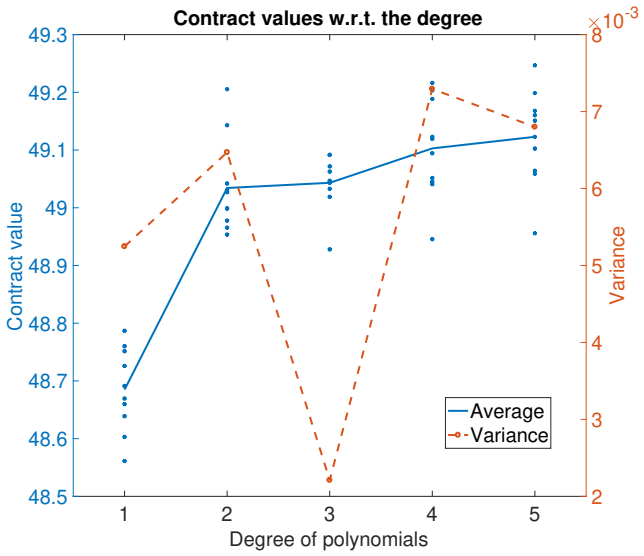
Note: The number of paths used: $G_1 = 100000, G_2 = 1000000$.

the other choices. Although the difference between values obtained by using **C.6** and **C.7** is relatively small, **C.7** remains the best explanatory vector because these two choices have the same computing effort for a fixed degree h, and **C.7** always gives better exercising rules. Indeed, this also evidences that the running average M explains the contract value better than the moving average A.

5.8.2 The degree of the complete set of polynomials

Since we have chosen **C.7**, $\mathcal{E} = (S, Z, M, I)$, as the explanatory vector, we now investigate how the degree of the CSP affects the contract values. Recall (5.29), since we have a fixed number of variables, $v = 4$, a high degree h means more basis functions in $\mathscr{P}_h^4$. The purpose of this subsection can therefore also be considered as investigating how the number of basis functions affects the contract value. Numbers of basis functions in $\mathscr{P}_h^4$ for different degrees h can be found in Table 5.4. We again simulate a pair of fixed sets of paths, one for the backward scheme and one for the forward scheme. Then we perform the LSMC on these two sets of paths by different degrees of the

FIGURE 5.3: Stability with respect to the degree h

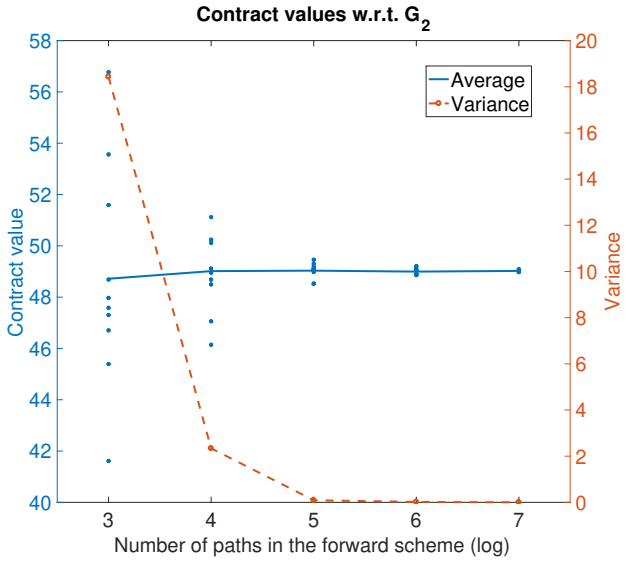


Note: The number of paths used: $G_1 = 100000, G_2 = 1000000$.

CSP. The results are shown in Figure 5.2. As we can see in Figure 5.2, higher degrees return higher contract values. The contract value grows rapidly from $h = 1$ to $h = 2$. Although the improvement of the contract value becomes smaller from $h = 2$ to $h = 4$, we can still observe a significant jump in Figure 5.2. In addition, the value obtained with $h = 5$ is only slightly higher than the value obtained with $h = 4$, thus demonstrating that the CSP of degree 4 is good enough to value such a GSA contract. Taking into consideration the efficiency, Figure 5.2 also shows the computing times with respect to the degree h. As we can see, the computing time grows almost exponentially with respect to the degree. The computing time is doubled from $h = 4$ to $h = 5$, and this gives us another reason to use $h = 4$. Also, the number of basis functions $\#(\mathscr{P}_h^4) = 70$ is reasonable since we have used a large number of simulated paths.

Next, we give the stability of the LSMC with respect to the degree in Figure 5.3. In Figure 5.3, we value the GSA contract ten times for each degree h. The smaller the variance of these ten valuations is, the more stability the

FIGURE 5.4: Stability w.r.t. G_2



NOTE: The number of paths in the backward scheme is $G_1 = 100000$, the degree of the CSP is $h = 2$.

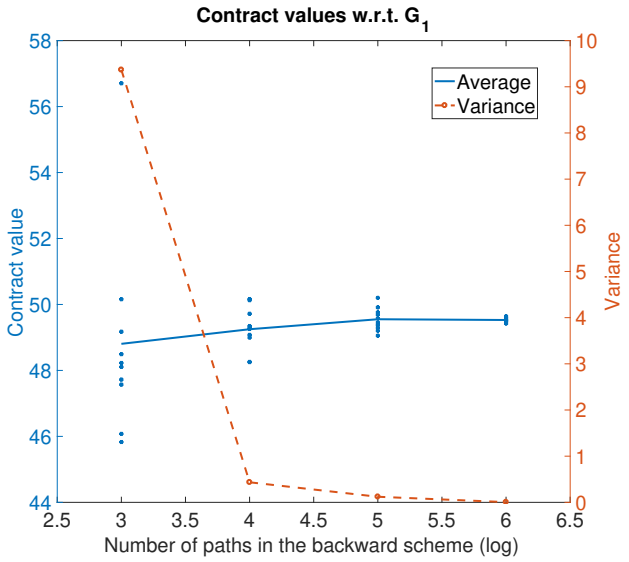
LSMC algorithm has. As we can see, the degree of CSP has no impact on the stability of the LSMC algorithm. In addition, the average value of these ten valuations increases as the degree increases. This also demonstrates that higher degrees give higher contract values.

5.8.3 The number of paths

In this section, we investigate the impact of the number of paths used in the LSMC algorithm on the valuation of the GSA contract.

The number of paths in the forward scheme First, we consider the number of paths, G_2, in the forward scheme. Figure 5.4 shows the stability of the LSMC algorithm with respect to the number of paths G_2. In Figure 5.4, we value the GSA contract ten times for each number of paths 10^n in the forward scheme, $n = 3, 4, 5, 6, 7$. As we can see, the variance of these ten valuations decreases dramatically as the number of paths increases. If we use $G_2 = 10^7$, all ten valuations give, more or less, the same value. In addition, the

FIGURE 5.5: Stability w.r.t. G_1

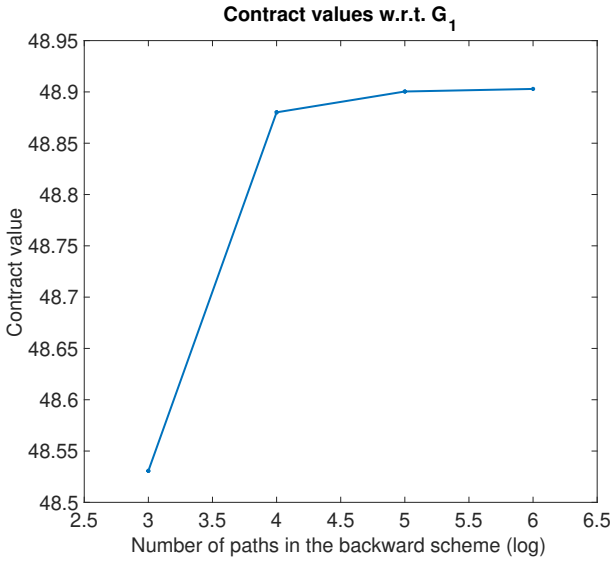


NOTE: The degree of the CSP is $h = 2$.

average of these ten valuations is quite stable from $G_2 = 10^3$ to $G_2 = 10^7$, demonstrating that increasing the number of paths in the forward scheme does not affect the contract value.

The number of paths in the backward scheme Although the value obtained through the backward scheme is highly biased (see Section 5.3), we can use this value to see the stability of the backward scheme. Figure 5.5 gives the results. In Figure 5.5, we again value the GSA contract ten times for each number of paths 10^n in the backward scheme, $n = 3, 4, 5, 6$. If the variance of these ten valuations is small, it means that all these valuations using the backward scheme give, more or less, the same exercising rule as that used in the forward scheme. As we can see, the larger G_1 is, the smaller the variance is. At $G_1 = 10^6$, it seems that the backward scheme already gives a good enough exercising rule since the variance is quite small. As will be seen later, however, we have to compromise to use fewer paths in the backward scheme than $G_1 = 10^6$ since the computational cost is too high. In addition, in Figure 5.5, the average of these ten valuations increases as G_1 increases. This

FIGURE 5.6: Contract values w.r.t. G_1

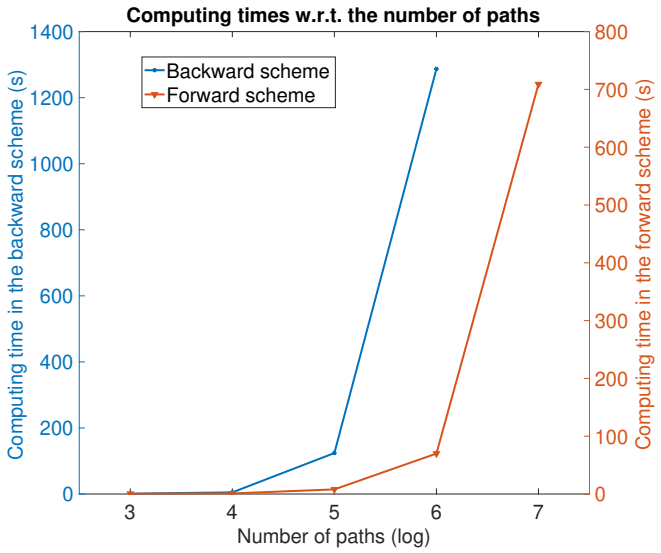


Note: The number of paths in the forward scheme is $G_2 = 1000000$. The degree of CSP is $h = 4$.

demonstrates that using more paths in the backward scheme gives a better exercising rule, as becomes clear in Figure 5.6. In Figure 5.6, for a fixed set of paths in the forward scheme, we apply the exercising rule obtained by using a different number of paths in backward scheme $G_1 = 10^n$, $n = 3, 4, 5, 6$. As we can see, the exercising rule obtained by using more paths in the backward scheme gives a higher value in the forward scheme. The improvement of the contract value is rather less significant from $n = 5$ to $n = 6$, indicating that, although the stability when using $G_1 = 10^5$ in the backward scheme is not very high (see Figure 5.5), it remains a reasonable choice since it gives a satisfactory contract value.

Computing time with respect to the number of paths Figure 5.7 gives the computing time with respect to the number of paths. For the backward scheme, we give the computing times when using $G_1 = 10^n$, $n = 3, 4, 5, 6$. For the forward scheme, we give the computing times when using $G_2 = 10^n$, $n = 3, 4, 5, 6, 7$. As we can see, if we use the same number of paths in both

FIGURE 5.7: Computing time w.r.t. the number of paths



Note: The degree of CSP is $h = 4$.

schemes, that is, $G_1 = G_2$, the backward scheme is much more time consuming than the forward scheme. This is because of the fact that we have to solve the minimization problem (5.6) in the backward scheme. At $G_1 = 10^6$, it already costs more than ten minutes to value even such a simple GSA contract. As we mentioned before, although the backward scheme is stable when using $G_1 = 10^6$, we may have to use fewer paths in the backward scheme, since the computing time will increase massively when dealing with a one-year contract ($N = 365$). The computing time of the forward scheme is also quite significant at $G_2 = 10^7$, which means we would also have to use fewer paths if we were attempt to value a one-year contract.

5.9 Conclusion

This chapter has proposed several simulation-based algorithms for pricing GSAs with indexation. Instead of using the value directly from the regression, we propose to use the LSMC algorithm with an exercising rule which

does not have the so-called anticipatory character. That is, the LSMC with an exercising rule does not assume knowledge of future contract values. An upper bound algorithm is also given and we have explained why this algorithm works on our GSA evaluation problem. We have proposed a benchmark algorithm which will be used in Chapter 6 to test the accuracy of the tree algorithm built in Chapter 4.

Furthermore, we have investigated which explanatory variables should be used in the LSMC algorithm. We have found that simply performing the regression on the gas price and the index cannot get a satisfactory result. That is, the contract value cannot be explained well by the gas price and the index. More explanatory variables should be used in the regression. Through numerical examples, we conclude that the combination of the gas price, the oil price, the running average price and the index explains the contract value better than other choices of explanatory variables. This gives about 4% higher contract values than the most commonly used explanatory variables: the gas price and the index. Our numerical examples also demonstrate that modelling the index through the running average variable is better than modelling it through the moving average variable.

Finally, we have investigated ways to get better outcomes though the LSMC algorithm. Using more basis functions generally gives better contract values, and using more paths in the backward scheme generally gives better exercising rules. Increasing the numbers of paths in both the backward scheme and the forward scheme can increase the stability of the LSMC algorithm. We also observe, however, that both increasing the number of basis functions and increasing the number of simulated paths can significantly increase the computing time.

Chapter 6

Implementation on GPUs

6.1 Introduction

Due to the high-dimensional nature of the GSA evaluation problem, it is quite challenging to implement the algorithms described in Chapter 4 in an efficient manner so that practitioners are able to use the approach to evaluate the real contracts. On the other hand, due to the nature of the tree methods, the algorithms we designed are highly parallelable, hence we can seek parallel computing platforms to implement our tree algorithm. In addition, when dealing with the evaluation of the GSA contract, the LSMC algorithm can also be very time consuming. As we mentioned in Remark 5.4 and Remark 5.5, the LSMC algorithm can also benefit from parallel computing. The amazing development in the graphics processing units (GPU) technology in recent years has provided us with a powerful tool to tackle such high-dimensional problems. Because of the many-core nature of the GPU, it has attracted a lot of attention in the area of computational finance. Works on the GPU implementation of simulation-based algorithms can be found in Abbas-Turki and Lapeyre (2009), Benguigui and Baude (2014), Cvetanoska and Stojanovski (2012), Dang, Christara and Jackson (2012), Fatica and Phillips (2013), Pagès and Wilbertz (2012), Leitao and Oosterlee (2015) and the references therein. Works on the GPU implementation of lattice-based algorithms can be found in Dang, Christara and Jackson (2009), Egloff (2012), Ganesan, Chamberlain and Buhler (2009), Jauvion and Nguyen (2008) and the references therein.

This chapter has two aims. First, we introduce GPUs and how they can be used for general computing purposes. Second, we implement both the tree algorithm and the LSMC algorithm and provide a comparison between the

two. This chapter is organized as follows: Section 6.2 and Section 6.3 give a brief introduction on the general-purpose computing on graphics processing units. Section 6.4 provides a description on how we can implement both the tree algorithm and the LSMC algorithm on GPUs. Section 6.6 provides numerical examples which analyze the performance of the tree algorithm and the LSMC algorithm. We draw conclusions in Section 6.7.

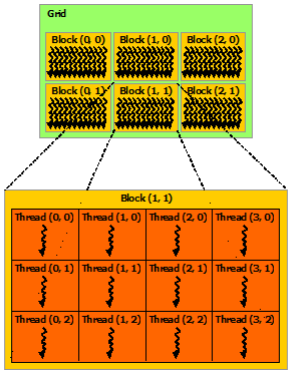
6.2 General-purpose computing on graphics processing units

General-purpose computing on graphics processing units (GPGPU) is the use of GPUs in fields not related to graphics processing. Typically, it means undertaking the computational work on the GPU instead of the CPU. One commonly used and successfully built general purpose parallel computing platform and programming model is the Compute Unified Device Architecture (CUDA), which was introduced by NVIDIA in 2006 (see NVIDIA (2018c)). CUDA can be treated as an extension of the traditional programming language C/C++, and hence eases the coding effort for programmers using C/C++. At the same time, other programming languages, such as FORTRAN, are supported. Mostly, CUDA can only be used on GPUs produced by NVIDIA itself. An alternative to CUDA is OpenCL (see Gohara, Shi and Stone (2010)) which can be used on both NVIDIA and non-NVIDIA GPUs. In this thesis, we use CUDA as our programming model. Since NVIDIA GPUs are widely installed on computers from personal laptops to high-performance supercomputers, it is not hard to find a CUDA-supported GPU.

6.3 The programming model

In order to present the implementation of the algorithm built in Chapter 4, we summarize the basics of the CUDA C/C++ programming model from a developer's point of view.

FIGURE 6.1: Thread hierarchy



Note: Reprint from NVIDIA (2018c).

6.3.1 Host, device and kernels

Although the reason we use CUDA is to perform parallel computing on the GPU, a CUDA programme is executed on both the CPU side and the GPU side. In fact, CUDA considers the GPU as a coprocessor to the CPU. In addition, CUDA uses its own terminology to identity these two sides: host refers to the CPU and its memory (host memory), device refers to the GPU and its memory (device memory). At runtime, the serial jobs are executed on the host while the parallel jobs are executed on the device. That is, the CUDA C/C++ code contains both the host code, written in C/C++, and the device code, written in C/C++, together with an extra declaration specifier. In fact, the host code and the device code are compiled separately. During compilation, the host code is forwarded to the C/C++ compiler (`gcc/g++` on Linux, `clang/clang++` on Mac and `cl.exe` on Windows). The device code is handled by the NVCC compiler (which is the compiler provide by NVIDIA, see NVIDIA (2018b)).

6.3.2 Thread Hierarchy

On the GPU, the thread is the smallest unit which performs computing. It is similar to the thread when we use open multi-processing (openMP) on a

multi-core CPU. The smallest unit of execution is called a warp which contains 32 threads. That is, even if we only want to launch a single thread, 32 threads of the hardware resource on the device will be used. Threads are grouped into a one-dimensional, two-dimensional or three dimensional block. Each block contains at most 1024 threads. Furthermore, blocks are grouped into a one-dimensional, two-dimensional or three-dimensional grid. Figure 6.1 shows the thread hierarchy. In the CUDA C/C++ programme, a kernel is a function which is called from the host and executed on the device. The kernel is defined with the declaration specifier `__global__`. A kernel must be called together with triple angle brackets `<<<dimGrid, dimBlock >>>`, where `dimGrid` and `dimBlock` are one-dimensional, two-dimensional or three-dimensional vectors which specify the size of the grid and block, respectively.

6.3.3 Memory Hierarchy

The device memory also has a hierarchy which contains several types of memory. Different memory types offer different characteristics and it is important to understand these in order to optimize the CUDA C/C++ program. Here we present some commonly used memory types.

Registers The register memory is analogous to the register on a CPU. It is thread local. That is, when launching a thread, this thread gets its own registers and these registers cannot be shared with other threads. The register memory has the same lifetime as the thread. Also, since the register is on-chip, it has the fastest speed. For the same reason, however, it is the smallest memory on the device.

Shared memory Each block has access to the shared memory and this memory is only shared among all threads in this single block. That is, if many blocks are launched at the same time, each of these blocks has its own shared memory. The shared memory has the same lifetime as the block. In addition, the shared memory is also on-chip and very fast.

Global memory The global memory on a GPU is analogous to the RAM on a CPU. When a kernel is called, all threads in all blocks can access the global

memory. The bandwidth[1] of the global memory is very large, but its speed is relatively slow since it is off-chip and has high latency. The global memory is mainly used to store large data which needs to be accessed from all threads.

Constant memory The constant memory can be considered as a part of the global memory but it is cached and offers faster reading speed than the global memory. Since the constant memory can only be written to from the host and it is read-only on the device, it is usually used to store constant parameters (such as the interest rate) which are intensively read from all threads.

Local memory When the registers for a thread are insufficient, the spilled data is stored in the local memory. The local memory can also be considered as a part of the global memory and has a slow speed.

In terms of the speed, the register is the fastest, then the shared memory and the constant memory while the global memory and the local memory are the slowest. The memory allocation is very important in CUDA, mainly because the reading and writing speed of the global memory and the local memory on the device is not as fast as the speed of the RAM on the host. There are two basic principles when programming in CUDA. First, try to make the kernels lightweight, since heavy calculations usually come with more register usage, and if the registers are insufficient, we have to use the local memory which causes the CUDA programme slow. Second, load the data which can be shared within a block into the shared memory from the global memory. This action reduces the use of the global memory and hence speeds up the programme.

6.3.4 The memory transfer

Besides the memory allocation, another key aspect in CUDA programming is the memory transfer. As mentioned above, the CUDA C/C++ code contains both the host code and the device code. At runtime, data has to be copied from the device to the host or from the host to the device so that the host code and the device code can use it, respectively. Since the host and the device are separate parts in a computer, the memory transfer between them is relatively

[1]The memory bandwidth is the maximum theoretical rate at which data can be read from or write into the memory.

slow, especially when transferring a lot of data. Although NVIDIA introduced the so-called unified memory, which can be accessed from both the host and the device, access to this memory still involves transferring data between both sides. Algorithms with lots of memory communications between the host and the device usually cost more computing time than those that do not. In addition, high-dimensional problems usually need more memory on the device to store data. When the memory is insufficient, one has to store data on the host. This leads to more communications between the host and the device and hence to increased computational time.

6.3.5 Our host and device

The GPU we work with is a GeForce GTX 1060 on an Acer Helios 300 laptop. Some details of this GPU are given below:

Number of CUDA cores:	1024
Total amount of global memory:	6078 MBytes
Total amount of constant memory:	65536 bytes
Total amount of shared memory per block:	49152 bytes
Total amount of registers per block:	65536

The CPU on the same laptop is an Intel(R) Core(TM) i7-7700HQ CPU @ 2.80 GHz together with 16 GB RAM.

6.4 Implementation on GPUs

In this section, we give an illustration of how we can implement our algorithm on GPUs. Recall that the evaluation of the GSA in the middle months is a five-dimensional problem (S, Z, M, I, Q) while the evaluations in the first month and the last month are four-dimensional (S, Z, M, Q) and three-dimensional problems (S, I, Q), respectively. In this chapter, we mainly focus on the evaluation in the middle months since it is this that requires the most computing effort. The evaluations in other months can be conducted in a similar

way, however. First recall the algorithm in the middle months:

$$V^l(S_{n,s}, Z_{n,z}, M_{n,m}, I_{n,i}, Q) = \max_{q\in[0,\bar{q}]} \Big\{ q \cdot (S_{n,s} - I_{n,i}) + \sum_{b,c=-1}^{1} p_{s,z,b,c} e^{-r\Delta t} \cdot V^l(S_{n+1,\mathfrak{g}(s,b)}, Z_{n+1,\mathfrak{h}(z,c)}, \mathcal{M}_{n,m,c}, I_{n,i}, Q+q) \Big\}. \tag{6.1}$$

Remark 6.1. Note that not everything is paralleled on the GPU. We only let the GPU do the heavy work, that is equation (6.1). Some easy work, such as the building of the two-dimensional trinomial tree, are implemented on the CPU. To ease the burden on the device, we can do some preparation before launching kernels on the device. Recall that, when $\mathcal{M}_{n,m,c}$ in (6.1) is not one of the values in the running average vector $\mathbf{M}_{n+1}$, we use linear interpolation (4.32). This involves finding $M_{n+1,\bar{m}}$ and $M_{n+1,\bar{m}+1}$. At any node (n, s, z) on the two-dimensional tree, the part

$$\frac{\mathcal{M}_{n,m,b} - M_{n+1,\bar{m}}}{M_{n+1,\bar{m}+1} - M_{n+1,\bar{m}}} \tag{6.2}$$

in (4.32) only depends on the value of the running average $M_{n,m}$ at time t_n and the oil price $Z_{n+1,\mathfrak{h}(z,c)}$, $c = -1, 1, 1$, at time t_{n+1} (see (4.18)). At time t_n, for each combination of $M_{n,m}$ and $Z_{n,z}$, we can find $\bar{m}$ and the value of (6.2) by following the up, middle and down movements on the oil tree before launching the kernels. In addition, at any node (n, s, z), for each combination of $S_{n,s}$ and $I_{n,i}$, we can also compute the instant payoff $(S_{n,s}-I_{n,i})$ in advance. The reason for these two actions is that, if we do these calculations inside a kernel, this kernel has to do the same calculations repeatedly, which wastes a lot of time.

Listing 6.1 gives the thread hierarchy of our CUDA programme. In Listing 6.1, `H` is the number of values in the running average vector $\mathbf{M}_n$, `Q` is the number of possible periods to date at time t_n, `N_s` and `N_z` are the total number of nodes on the gas tree and the oil tree at time t_n, respectively. `dim3` is an integer vector type in CUDA which defines the dimensions of the block and the grid.

LISTING 6.1: Block and grid dimensions

```
dim3 dimBlock(H);
dim3 dimGrid(N_s, N_z, Q);
```

Listing 6.2 gives the code on how we launch our kernel. The kernel which implements our algorithm (6.1) is `GAS_Middle_Months()`.

LISTING 6.2: Call the kernel

```
for(int i = 0; i < H; i++)
{
   GAS_Middle_Months<<<dimGrid, dimBlock>>>();
}
```

Listing 6.1 and Listing 6.2 mean that, if we have a strong enough device, at any time t_n, for each index $I_{n,i}$ in the index vector $\mathbf{I}_n$, the calculations of the values $V^l(S_{n,s}, Z_{n,z}, M_{n,m}, I_{n,i}, Q)$ for all combinations of $S_{n,s}$, $Z_{n,z}$, $M_{n,m}$ and Q are paralleled on the device. Listing 6.3 shows how we use shared memory to accelerate the reading and writing speed on the device.

LISTING 6.3: Shared memory allocation

```
__shared__ float linp_m[H][3];
__shared__ int m_bar[H][3];
__shared__ int n_1[6];
__shared__ float prob[9];
__shared__ float i_p;
__shared__ float n_1_value[3][3][H];
```

In Listing 6.3, `linp_m[H][3]` stores the values of (6.2), `m_bar[H][3]` stores the values of $\bar{m}$, `n_1[6]` stores $\mathfrak{g}(s, b)$ and $\mathfrak{h}(z, c)$, `prob[9]` stores the probabilities associated with the nine possible movements on the two-dimensional tree, `i_p` stores the instant payoff and `n_1_value[3][3][H]` stores the value of the GSA at time t_{n+1}. As we can see, `n_1[6]`, `prob[9]` and `i_p` consume only a very small amount of memory. `linp_m[H][3]`, `m_bar[H][3]` and `n_1_value[3][3][H]` may consume more memory since the sizes of these arrays depend on an integer `H`, which is the number of values in the running average vector $\mathbf{M}_n$. If we increase `H` to get a better approximation of the running average M, these arrays consume more memory, and this may reduce the number of blocks which can be paralleled at the same time. The

advantage of the memory allocation in Listing 6.3, however, is that, if we increase `H`, the number of threads in each block is also increased. That is, even with less blocks paralleled, we still have a lot of threads running at the same time.

Remark 6.2. In each month, we evaluate the GSA for each value of index in the index vector **I**. It means that, for each `i` in Listing 6.2, the kernel `GAS_Middle_Months` can also be paralleled. We do not do that mainly because of the shortage of memory on the device. Although the global memory is the largest memory on the device, its size is usually smaller than the RAM on the host. For instance, the size of the global memory on our device (see Section 6.3.5) is 6 GB while the size of the RAM on our host is 16 GB. Even when `i` in Listing 6.2 is not paralleled, we show in Section 6.6 that the computing time is satisfactory.

Remark 6.3. As we mentioned in Section 6.3.4, the data transfer between the host and the device can significantly slow the CUDA programme. In CUDA programming, memory transfers usually happen when the data is too large to store on the device. The advantage of our algorithm is that, under normal circumstances, we only need to do the data transfer when we want to retrieve specific data from the device. The whole evaluation of the GSA can be done on the device. This because of the matching point condition (4.35). At the matching point between two consecutive months, we only need to store data when the oil price equals the value of the running average, hence, one dimension has been reduced.

6.5 LSMC on GPUs

The LSMC algorithm can also be performed on GPUs. The most inefficient part of the LSMC algorithm is the backward scheme (see Figure 5.7). This is because the backward scheme involves solving the minimization problem (5.6). Obviously, this minimization problem cannot be paralleled since it requires information from all paths. In Fatica and Phillips (2013), the authors proposed a method to price American options on GPUs, which also involves solving a minimization problem. In this section, we apply their method to our problem. Recall notations in Remark 5.5, the minimization problem can

be written in the following form:

$$\boldsymbol{\phi}_n \boldsymbol{\beta}_n^{Q,q} = \boldsymbol{V}_n^q, \tag{6.3}$$

where $\boldsymbol{\phi}_n$ is a $G_1 \times K_b$ matrix, $\boldsymbol{\beta}_n^{Q,q}$ is a $K_b \times 1$ vector and $\boldsymbol{V}_n$ is a $G_1 \times 1$ vector. From (6.3), it can be easily seen that

$$\left(\boldsymbol{\phi}_n^{\text{trans}} \boldsymbol{\phi}_n\right) \boldsymbol{\beta}_n^{Q,q} = \left(\boldsymbol{\phi}_n^{\text{trans}} \boldsymbol{V}_n^q\right), \tag{6.4}$$

where $\boldsymbol{\phi}_n^{\text{trans}}$ is the transpose of $\boldsymbol{\phi}_n$. Now, we can find the optimal coefficient vector by solving (6.4). Since $\boldsymbol{\phi}_n^{\text{trans}} \boldsymbol{\phi}_n$ is a $K_b \times K_b$ matrix and $\boldsymbol{\phi}_n^{\text{trans}} \boldsymbol{V}_n^q$ is a $K_b \times 1$ vector, (6.4) is easy and fast to solve. Since solving (6.4) also requires all the information in $\boldsymbol{\phi}_n^{\text{trans}} \boldsymbol{\phi}_n$ and $\boldsymbol{\phi}_n^{\text{trans}} \boldsymbol{V}_n^q$, it can only be done on a single core (or thread). The matrix multiplication $\boldsymbol{\phi}_n^{\text{trans}} \boldsymbol{\phi}_n$ and $\boldsymbol{\phi}_n^{\text{trans}} \boldsymbol{V}_n^q$ can by paralleled on GPUs, however. The matrix multiplication is a classic CUDA programming problem, NVIDIA also provides the CUBLAS library (see NVIDIA (2018a)) to ease the programming effort. In addition, since all the paths in both the backward scheme and the forward scheme are generated independently, the generations of all paths can be paralleled on the GPU. This is implemented in a similar manner as the paths generation on the CPU, except we assign one path to one thread on the GPU and then move forward in time to generate the full path.

However, computing $\boldsymbol{\beta}_n^{Q,q}$ through (6.4) has a downside. Although (6.4) is theoretically correct, it causes problems when implemented on a computer. Recall that each row in the matrix $\boldsymbol{\phi}_n$ contains the values of the basis functions with respect to the realization of the explanatory vector $\mathcal{E}$. Recall notations in (5.18), after the matrix multiplication, the entry in the k_1th row and k_2th column of the matrix $\boldsymbol{\phi}_n^{\text{trans}} \boldsymbol{\phi}_n$ is given by

$$\sum_{g=1}^{G_1} \phi_{k_1}(\mathcal{E}_n^{(g)}) \phi_{k_2}(\mathcal{E}_n^{(g)}), \tag{6.5}$$

$k_1,\ k_2 = 1, \ldots, K_b$. When we use many simulated paths, a not very high degree h of the complete set of polynomials gives a very high value of (6.5). The information of all simulated paths at time t_n is suppressed into (6.5). Due to the storage limit, some information is lost. We demonstrate the effect

TABLE 6.1: Parameters set for numerical examples

$\alpha_S = 5$	$\sigma_S = 0.5$	$\alpha_Z = 6$	$\sigma_Z = 0.6$	$r = 0.05$
$\rho = 0.5$	$q_{\min} = 0$	$q_{\max} = 1$	$MB = 270$	$ACQ = 360$
$T = 1$	$N = 360$	$D = 30$	$L = 12$	

of this problem in Section 6.6.

Remark 6.4. Recall Remark 5.4 and Example 5.1, when we implement the LSMC algorithm on the CPU, we can apply (5.16) to save a lot of time. The advantage of using (5.16) weakens if we have many parallel cores. This issue does not concern the CPU since the CPU usually contains four or eight cores. When we implement the LSMC algorithm on the GPU, we have hundreds of cores available. However, as we mentioned before, we mainly use cores on the GPU to compute the matrix multiplications $\boldsymbol{\phi}_n^{\text{trans}}\boldsymbol{\phi}_n$ and $\boldsymbol{\phi}_n^{\text{trans}}\boldsymbol{V}_n^q$ in (6.4). The matrix multiplications for different periods to date Q are not paralleled due to the limited storage on the GPU. In addition, (6.4) is solved by using a single core. Even using a single core, (6.4) can be solved very fast, since the sizes of $\boldsymbol{\phi}_n^{\text{trans}}\boldsymbol{\phi}_n$ and $\boldsymbol{\phi}_n^{\text{trans}}\boldsymbol{V}_n^q$ are small. Hence, the advantage of using (5.16) is much less significant on the GPU compared with it on the CPU. Nonetheless, we can always use (5.16) to save some time.

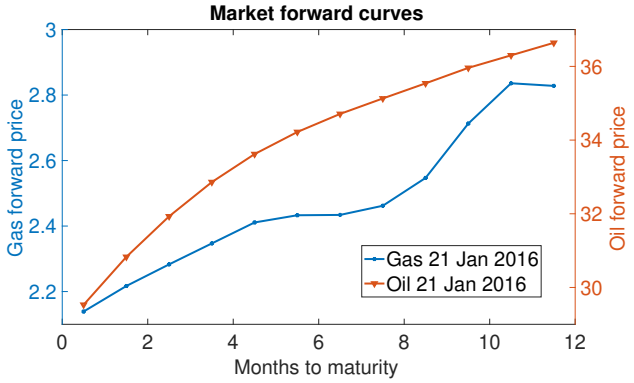
6.6 Numerical examples

In this section, we provide numerical examples which evaluate the GSAs by using the different algorithms we have built in this thesis. These algorithms are the tree algorithm (see Chapter 4) and the LSMC algorithm using an exercising rule (see Algorithm II). When not variable, we use parameter values in Table 6.1. We use both flat forward curves and real forward curves in our numerical examples. We assume the flat forward curves are given by

$$F^S(0,t) = 100, \quad F^Z(0,t) = 100 \tag{6.6}$$

for $t \in [0,T]$. The real forward curves are shown in Figure 6.2. For the sake of the comparison between examples, we scale both the gas forward curve and the oil forward curve in Figure 6.2 such that $F^S(0,0) = F^Z(0,0) = 100$.

FIGURE 6.2: Market forward curves



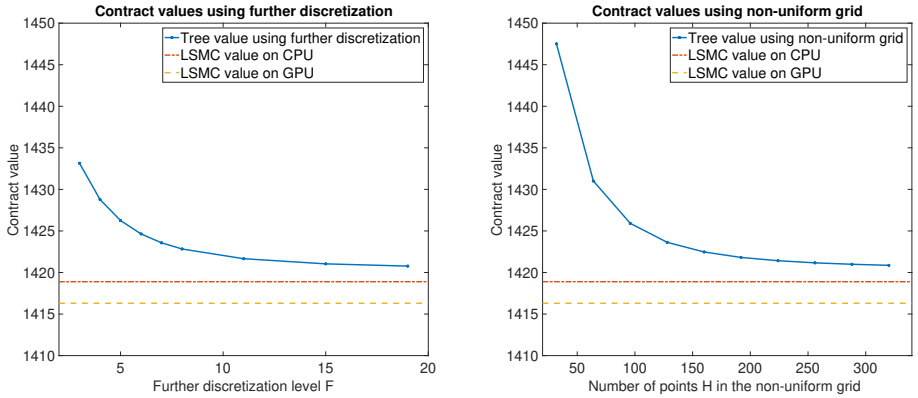
In terms of the non-uniform grid, when not variable, we let the constants $w_n = \frac{1}{3}$, $n = 1, 2, \ldots, N$, which control the fraction of points $M_{n,m}$ that lies in the neighbourhood of M_n (see (4.22)). Unless indicated otherwise, we use the contract values obtained by using the LSMC algorithm as pre-computed benchmark values. The LSMC algorithm has different speeds with GPUs and CPUs. In addition, unless indicated otherwise, the computing time appearing in different figures only refers to the tree algorithm.

6.6.1 Contract values

In this subsection, we investigate the performance of our algorithms in terms of the contract values. In the rest of this section, the tree FD values refer to the contract values obtained by the tree algorithm with further discretization (4.19), while the tree NUG values refer to the contract values obtained by the tree algorithm with the non-uniform grid (4.22), and the LSMC values refer to the contract values obtained by the LSMC algorithm using an exercising rule (see Algorithm II).

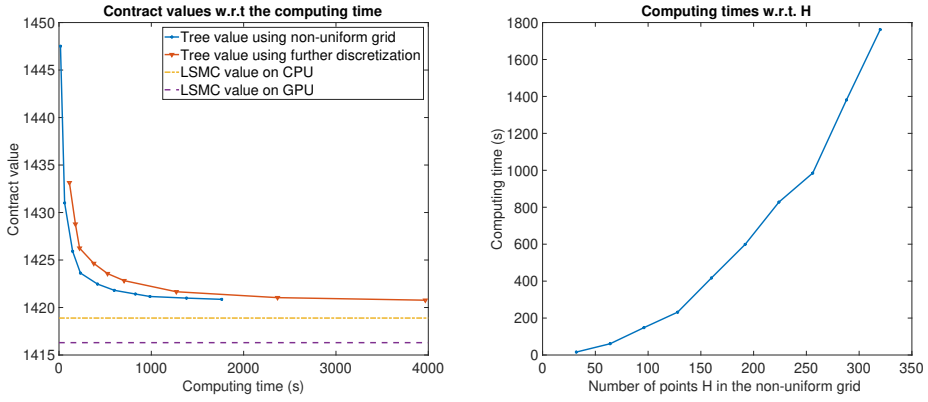
First, we evaluate the GSA contract with the flat forward curves (6.6). Figure 6.3 gives the comparison of contract values obtained by using different algorithms. In the left panel of Figure 6.3, F is the positive integer we use to do the further discretization (see (4.19)). A larger value of F results in a smaller value of ΔM, and this further leads to a denser grid in the running average direction. In the right panel of Figure 6.3, H is the number of values in the

FIGURE 6.3: Contract values w.r.t. the average discretization



Note: The standard error of the LSMC values on the CPU and on the GPU are 2.317 and 2.332, respectively. The upper bound obtained by Algorithm III is 1611.71 where 100000 paths have been used.

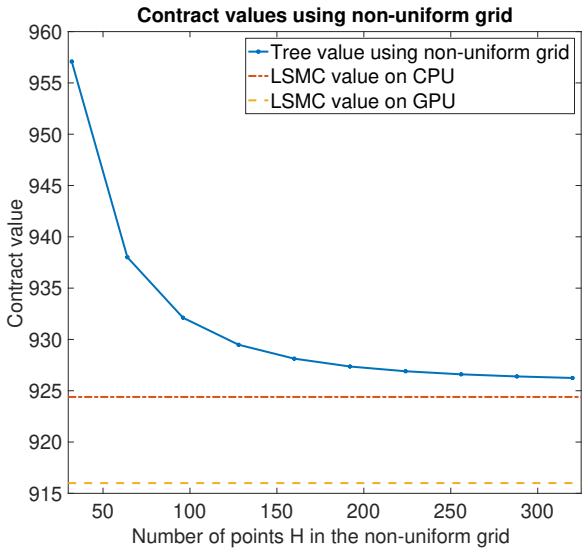
FIGURE 6.4: The convergence and the computing time



running average vector $\mathbf{M}$ and a larger value of H also gives a denser grid in the running average direction. By (4.23), a denser grid in the running average direction means a denser grid in the index direction. As we can see, both the tree FD value and the tree NUG value converge to but remain above the LSMC values. Furthermore, it seems that the tree FD value converges faster to the LSMC values than the tree NUG value. By (4.20), however, the running average vector $\mathbf{M}$ already contains 73 values at $F = 3$. Recall Theorem 4.2, by using the further discretization, we obtain the first order consistency when $F = \mathcal{O}(\Delta t^{-\frac{1}{2}})$. In our numerical example, we can achieve the first order consistency at $F = 19$. At $F = 19$, the number of values in the running average vector $\mathbf{M}$ has reached 457. As we can see in Figure 6.3, the tree NUG value at $H = 320$ already achieved the same value as the tree FD value at $F = 19$. This becomes clear in the left plot of Figure 6.4, which shows the contract values with respect to the computing time by using these two discretization method. As we can see, the tree NUG value converges much faster to the LSMC values than the tree FD value. This demonstrates that the non-uniform grid is a better discretization method than the further discretization. In the rest of this chapter, therefore, when we implement the tree algorithm, we only use the non-uniform grid.

In terms of the LSMC values, as we can see in Figure 6.3, the LSMC value on the GPU is below the LSMC value on the CPU. This is mainly because of the problem we mentioned in Section 6.5. That is, when we apply (6.4), some information of the simulated paths is lost while implementing the LSMC algorithm on a computer. However, the LSMC on the GPU has a great advantage: it is much more efficient than the LSMC on the CPU. In our numerical example, with 100000 paths in the backward scheme, 1000000 paths in the forward scheme, and the complete set of polynomials of degree 4, the LSMC on the CPU takes nearly 20 hours to price our GSA contract. Even if we parallel the LSMC algorithm on four cores (the CPU we use contains four cores), it still needs more than five hours. When the LSMC is implemented on the GPU, however, only 244 seconds are needed. Recall Remark 4.5, when we have a larger H, the computational cost of the tree algorithm is higher. The computing time with respect to H is shown in the right plot of Figure 6.4. As we can see, the computing time grows rapidly as H increases. Compared with the LSMC on the CPU, however, the tree algorithm is still very fast.

FIGURE 6.5: Contract values by using real forward curves

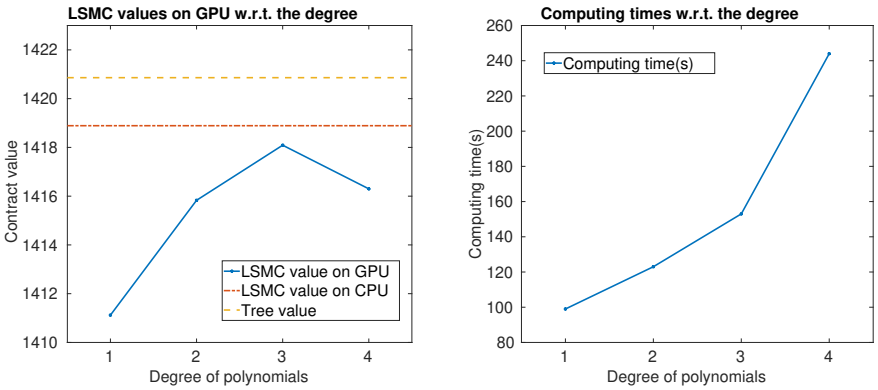


Note: The standard errors of the LSMC values on the CPU and GPU are 2.543 and 2.561, respectively. The upper bound obtained by Algorithm III is 1172.02 with 100000 paths used.

Now, we also give an example of pricing our GSA using real forward curves (see Figure 6.2). Figure 6.5 gives the comparison of contract values obtained by using different algorithms. We have omitted the tree FD value, since the tree NUG value outperforms the tree FD value, as we investigated before. As we can see, Figure 6.5 preserves the same characteristics as the right plot of Figure 6.3.

Another thing worth mentioning is that, no matter whether we use flat forward curves or real forward curves, the contract values obtained by both the tree algorithm and the LSMC algorithm are below the upper bound obtained by Algorithm III (see notes of Figure 6.3 and Figure 6.5). On the one hand, it demonstrates that our tree algorithm and LSMC algorithm are producing reasonable results in a reasonable time frame. On the other hand, it also demonstrates that Algorithm III overestimates the contract by about 14% when using flat forward curves and by about 25% when using real forward curves because of the perfect foresight.

FIGURE 6.6: The LSMC value on GPU



Note: 100000 paths and 1000000 paths have been used in the backward scheme and the forward scheme, respectively. From the degree $h = 1$ to $h = 4$, the standard errors of the LSMC values on the GPU are 2.324, 2.319, 2.321 and 2.561, respectively. The standard error of the LSMC value on the CPU with the degree $h = 4$ is 2.543. The tree value is obtained by using non-uniform grid with the number of points $H = 320$. In addition, the computing time in this figure refers to the LSMC algorithm on the GPU.

6.6.2 The LSMC algorithm on the GPU

In Figure 6.4, we see that the LSMC value on the GPU is less optimal compared with the LSMC value on the CPU. In this subsection, we seek a way to get a better contract value by using the LSMC algorithm on the GPU. Recall Section 6.5, when we apply (6.4) with many simulated paths and a high degree of the complete set of polynomials, (6.5) gives a very large value. Due to the storage limit, some information is lost. There are three ways to overcome this issue, we can use less simulated paths, use a lower degree of the complete set of polynomials, or do both. The Monte Carlo algorithms are based on convergence of the sampled average to the mathematical one, this is the reason why we can get better and more stable contract values by using larger sets of simulated paths (see Figures 5.4, 5.5 and 5.6). That is, as investigated in Section 5.8.3, we can get very different contract values if we do not use sufficient large sets of paths in the LSMC algorithm. Therefore, we seek to

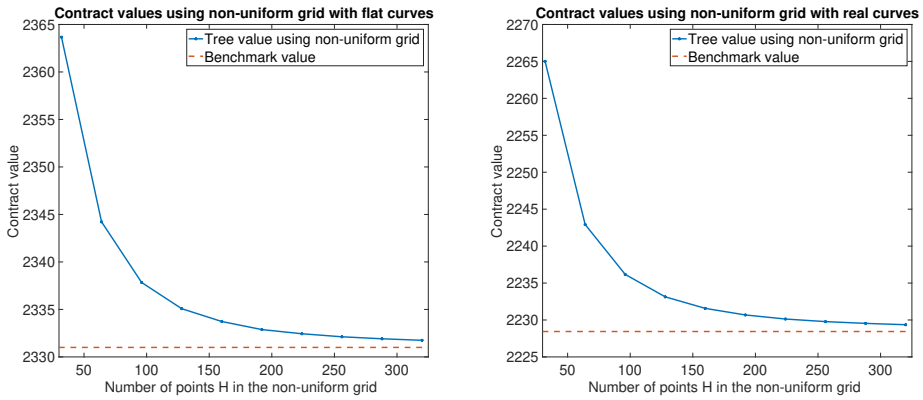
use a lower degree of the complete set of polynomials instead of decreasing the number of simulated paths. Figure 6.6 reports the outcome. As we can see in the left panel of Figure 6.6, from the degree of 1 to the degree of 3, the contract value increases as the degree of polynomials increases, which is consistent with the observation in Figure 5.2. Figure 5.2 also indicates that, if we further increase the degree of polynomials, we should get a higher contract value. However, the left panel of Figure 6.6 shows that the contract value at the degree of 4 is even lower than the contract value at the degree of 3. This is because that, at this point, the LSMC on the GPU starts to suffer from the problem we mentioned in Section 6.5. At the degree of 3, the LSMC value on the GPU is only slightly lower than the LSMC value on the CPU with the degree of 4. The difference is mainly because the LSMC algorithm on the CPU dose not suffer from the loss of information and can get a better exercising rule by using the complete set of polynomials of degree 4 (see Section 5.8.3). The right panel of Figure 6.6 shows the computing time with respect to the degree of polynomials. At the degree of 3, the LSMC algorithm on the GPU only requires 153 seconds to evaluate such a contract. That is, we can use a lower degree of polynomials to overcome the problem we mentioned in Section 6.5, and, at the same time, achieve a faster implementation.

6.6.3 The accuracy of the tree algorithm

Conparison between the tree algorithm and Algorithm IV Now, we investigate the accuracy of our tree algorithm by a comparison between the tree algorithm and Algorithm IV. That is, we let the minimum bill MB in Table 6.1 be 0, meaning that there is no penalty involved in our GSA. Figure 6.7 reports the results. At $H = 320$, the tree value with flat forward curves is only 0.03% above the benchmark value, while the tree value with real forward curves is only 0.04% above the benchmark value. Since 10 million paths have been used in Algorithm IV, these two benchmark values should be trustworthy.

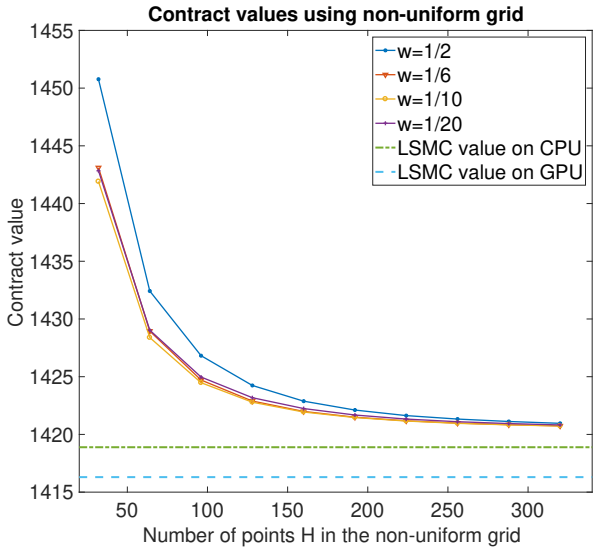
The impact of w_n Next, we investigate the impact of w_n on our tree algorithm. Recall that w_n, $n = 1, 2, \ldots, N$, is the constant which controls the fraction of points $M_{n,m}$ in the running average vector $\mathbf{M}_n$ that lie in the neighbourhood of M_n. In addition, a smaller w_n gives more points in the neighbourhood of M_n. In our examples, we let w_n be a constant w for all

FIGURE 6.7: Contract values when penalties are not involved



Note: The standard errors of the benchmark value with flat curves and the benchmark value with real curves are 0.507 and 0.51, respectively. 10 million paths have been used in Algorithm IV.

FIGURE 6.8: Contract values w.r.t. w

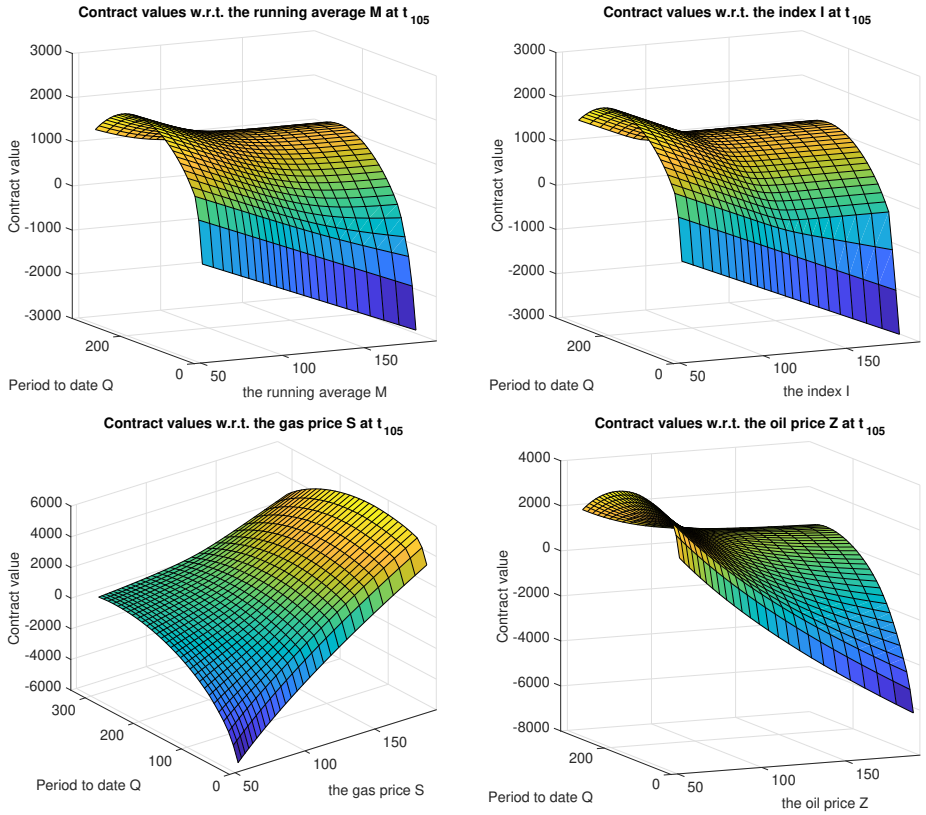


$n = 1, 2, \ldots, N$. Figure 6.8 reports the impact of w on our tree algorithm when pricing the GSA contract with flat forward curves. As we can see, both the convergence speed when using $w = \frac{1}{6}$ and the convergence speed when using $w = \frac{1}{10}$ outperforms the convergence speed when using $w = \frac{1}{2}$. That is, when more points in the non-uniform grid lie in the neighbourhood of M_n, the convergence speed of our tree algorithm is faster. This is quite understandable. Recall that M_n equals the value of the running average M by assuming that the oil price moves on the oil tree by following the middle branches only (see (4.21)). Due to the mean-reverting nature of the oil price, the oil price is unlikely to be too high or too low for a long time period. It follows that the value of the running average is even more unlikely to be too large or too small. Indeed, this is the reason why the indexation of the GSA exists: to smooth undesired volatility effects. In addition, as we can observe in Figure 6.8, the performance of our tree algorithm is worse when using $w = \frac{1}{20}$ than when using $w = \frac{1}{10}$. This means that, although the movement of the running average is relatively stable, assigning too many points in the neighbourhood of M_n will not benefit the tree algorithm. No matter which value of w we have used in Figure 6.8, however, the tree values of different values of w all converge to, more or less, the same value, at $H = 320$. This demonstrates that, when there are a lot of points in the non-uniform grid, the tree algorithm is unlikely to be affected by the value of w. Figure 6.8 also demonstrates that, if the non-uniform grid contains fewer points, we can use a smaller w to get a better contract value. When using $w = \frac{1}{10}$, the contract value has only changed 0.08% from $H = 160$ to $H = 320$. At the same time, the computing time at $H = 160$ is only 417 seconds, which is very competitive with the LSMC on the GPU. Even with $w = \frac{1}{2}$, the contract value has only changed by 0.13% from $H = 160$ to $H = 320$. This demonstrates that we can still get a satisfactory contract value with fewer points in the non-uniform grid, even with a relatively large w.

6.6.4 Value surfaces and decision surfaces from the tree algorithm

In this section, we investigate the value surfaces and the decision surfaces by evaluating the GSA with flat forward curves using the tree algorithm.

FIGURE 6.9: Value sufaces at time t_{105} by the tree algorithm



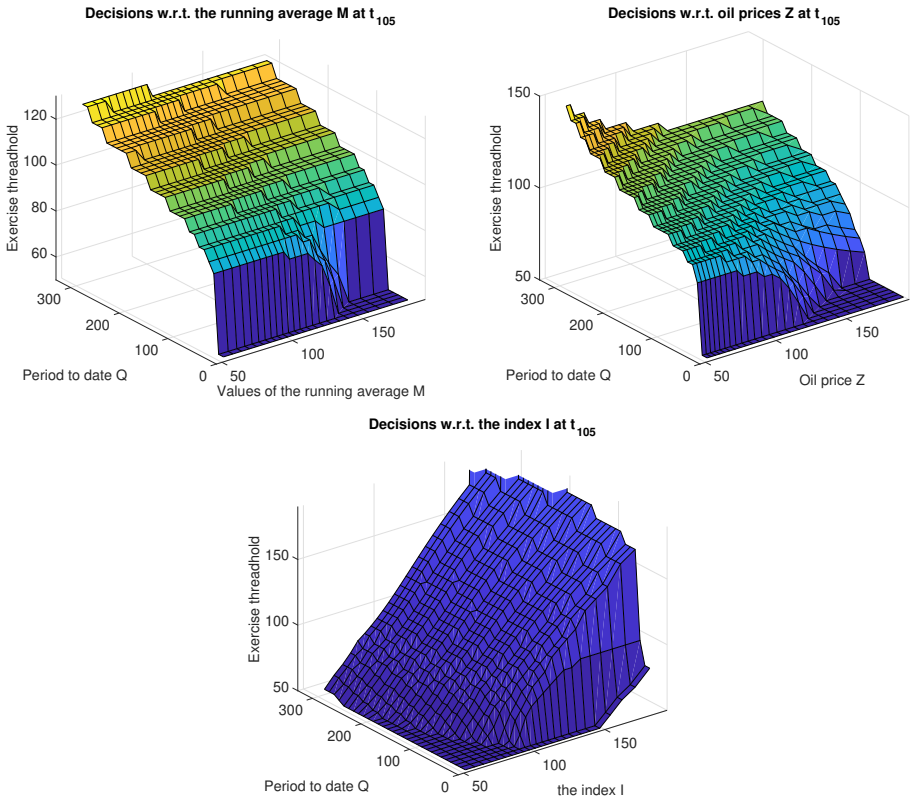
Note: When not variable, these surfaces are drawn where the gas price is 98.815, the oil price is 98.543, the value of the running average is 108.835 and the index is 108.909.

Value surfaces Figure 6.9 shows the value surfaces at time t_{105}, which is the 15th day of Month 3. As we can see, the value surfaces with respect to the running average, the index and the oil price have similar patterns. This is due to the fact that both the running average and the oil price influence the contract values by affecting the index in the coming month. Also, the contract value is more sensitive to the oil price than the running average and the index, especially when the oil price is low. This is because, if the current oil price is low, then the oil price in the rest of this month is possibly in the low price regime. This leads to a low index in the coming month. In addition, in

the surfaces with respect to the running average, the index and the oil price, for a fixed period to date Q, the contract value is stable when these corresponding values are high. This is because the running average, the index and the oil price are large enough to make sure the buyers avoid exercising gas rights under the contract. The bottom-left surface of Figure 6.9 shows that the contract value increases as the gas price increases. This is because the buyers are willing to take gas under the GSA contract when the gas price is high. This action not only gives profits to the buyers from the instant payoffs, but also reduces or possibly avoids the penalties since more gas has been taken under the contract.

Decision surfaces Recall that, we have defined the exercise threshold as the gas price equal to or above which the buyer would be better to take the daily maximum $q_{\max}$ in Chapter 3. A low exercise threshold means that the buyer is willing to take gas under the contract, while a high exercise threshold means the buyer is reluctant to take gas. For the buyers, the most important thing is to make the optimal daily exercise decision at the beginning of each gas day, based on the current gas price, oil price, index price and the value of the running average. We present decision surfaces obtained by using the tree algorithm at time t_{105} in Figure 6.10. As we can see in these surfaces, there is a big jump in the exercise threshold when the period to date is 15. This is due to the requirement to meet the minimum bill. At day 105, there are $360 - 105 = 255$ exercise opportunities left before the end of the contract. If the period to date is 15 at day 105, then the buyer can only avoid the penalty at the end of the contract by taking the daily maximum $q_{\max}$ on the remaining days. That is, the buyer has to take gas under the contract even when the instant payoff is negative. This feature is more clear in Figure 6.11 which shows the decision surfaces at t_{225}. At day 225, there are $360 - 225 = 135$ exercise opportunities left before the end of the contract. If the period to date is 135 at day 225, then the buyer can only avoid the penalty by taking the daily maximum $q_{\max}$ on the remaining days. As we can see in the top-right plot of Figure 6.11, the exercise threshold is higher when the index is very large. This is because the index is so large that the buyer loses more profits from the instant payoff than the penalty. Again, as we can see in the top two surfaces of Figure 6.10, the exercise threshold is more sensitive with

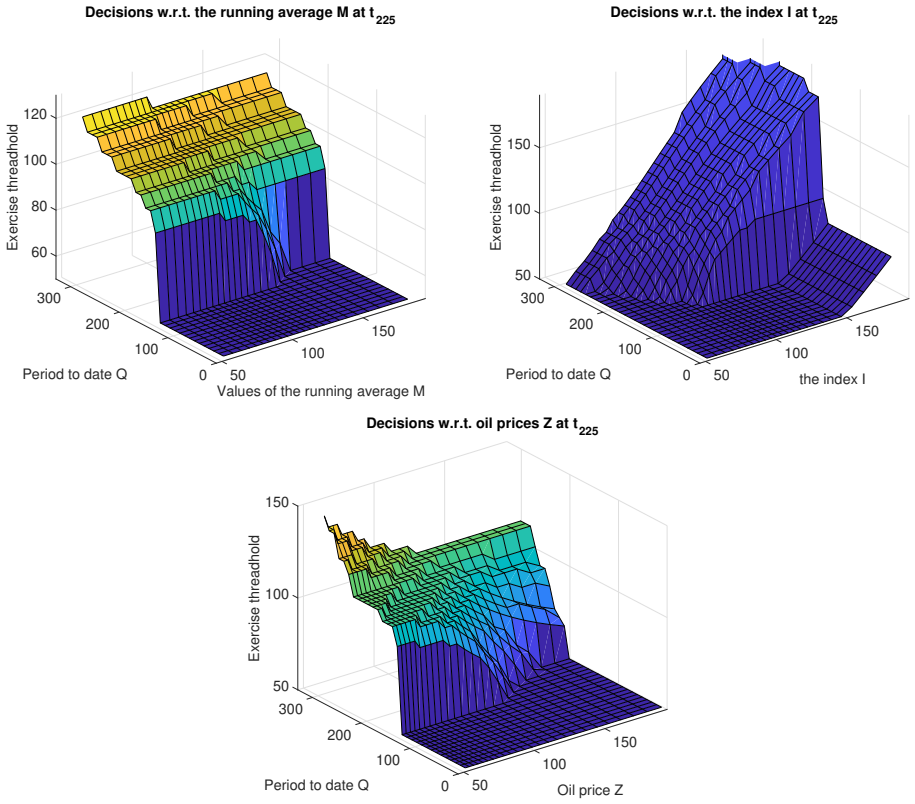
FIGURE 6.10: Decision sufaces at time t_{105} by the tree algorithm



Note: When not variable, these surfaces are drawn where the oil price is 98.543, the value of the running average is 108.835 and the index is 108.909.

respect to the oil price than the running average. In addition, the decision surfaces with respect to the running average and the oil price have different patterns with the surface with respect to the index. The exercise threshold increases as the running average and the oil price decrease. This is because, if the current oil price and the running average is low, then the index in the coming month is possibly very small, hence the buyers are willing to save exercise opportunities in order to take advantages of the small index in the coming month. The exercise threshold increases as the index increases. This is because the buyers do not benefit from a high index.

FIGURE 6.11: Decision sufaces at time t_{225}

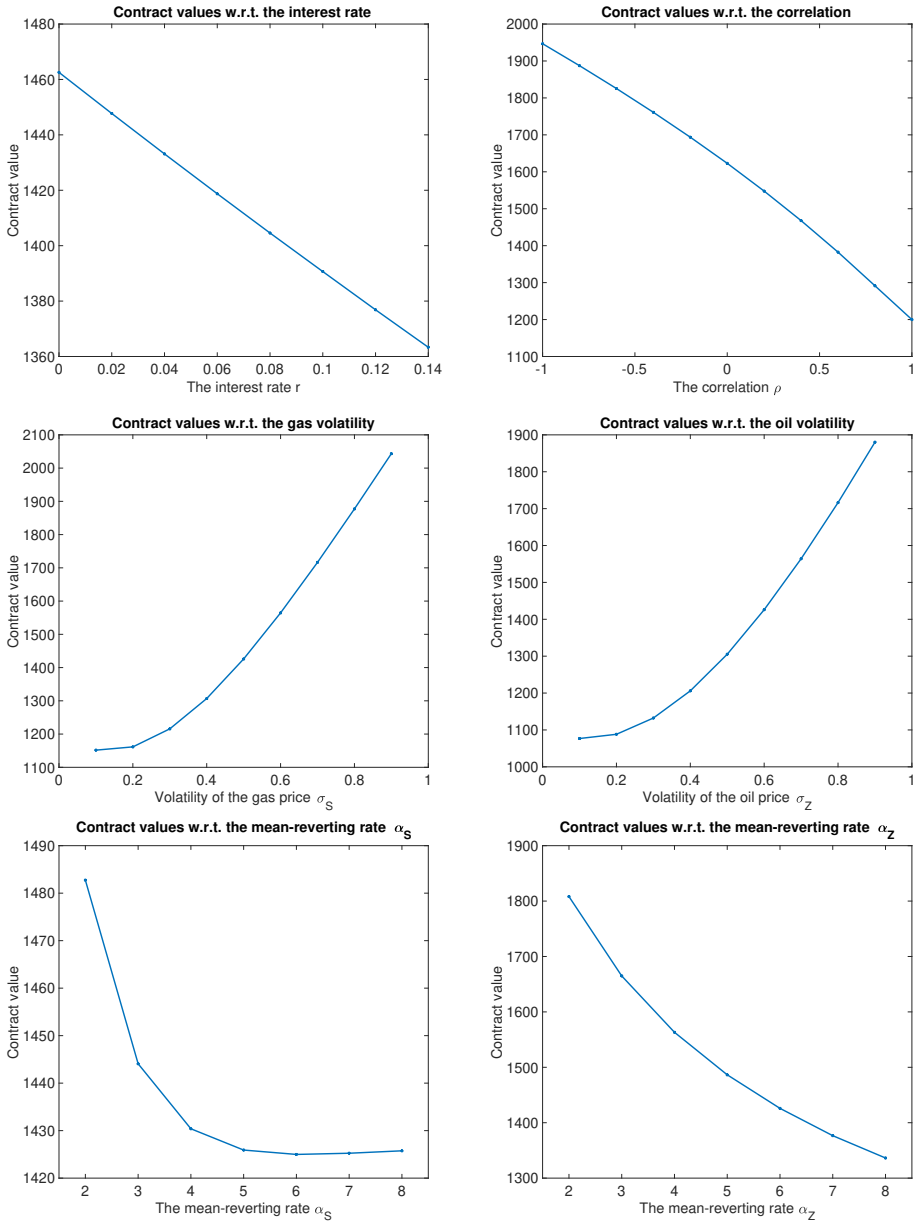


Note: When not variable, these surfaces are drawn where the oil price is 98.5, the value of the running average is 108.774 and the index is 108.775.

6.6.5 Contract values with respect to parameters

Due to the nature of the LSMC algorithm, for the same GSA contract, separate valuations can give different contract values (see the stability test in Section 5.8). On the contrary, for the same GSA contract, with a fixed H and w, the tree algorithm gives the same contract value through separate valuations. Hence, the tree algorithm provides a powerful tool to analyze how the parameters of a GSA contract affect the contract value. In this subsection, we investigate the impact of these parameters by using the tree algorithm. Again, we focus on the GSA with flat forward curves. Figure 6.12 shows how

FIGURE 6.12: Values with respect to parameters



the contract value changes when the values of parameters change. As we can see, the contract value decreases when the interest rate increases, since a high interest rate reduces the value of future cashflows. Also, the contract value decreases as the correlation increases. This is because the buyers have less opportunities to make high profits when the gas price and the oil price move in the same direction. In addition, the contract value increases in both the volatility of the gas price and the volatility of the oil price. This is because the buyers could implement a more flexible trading strategy when the gas price or the oil price is fluctuating more. Another thing worth mentioning is that the volatility of the gas price contributes more to the contract value than the volatility of the oil price. This is because the volatility of the oil price is smoothed under the indexation principle. Furthermore, generally speaking, the contract value decreases in both the mean-reverting rates of the gas price and the oil price. As we can see in the bottom-left plot of Figure 6.12, the contract value decreases rapidly from $\alpha_S = 2$ to $\alpha_S = 5$, before becoming quite stable as the mean-reverting rate of the gas price increases. This demonstrates that α_S loses its impact on the contract if it is large enough. As we can see in the bottom-right plot of Figure 6.12, compared with α_S, the mean-reverting rate of the oil price α_Z has a much greater impact on the contract value. This is because, when we have a small α_Z, the value of the index can be quite different in different months. This gives the buyers opportunities to take advantage of a low index and avoid taking gas under a high index. Recall Remark 4.5, different mean-reverting rates give different sizes of the trinomial tree, hence, the computing time of the tree algorithm should be related to α_S and α_Z. Figure 6.13 reports the computing time with respect to these two mean-reverting rates.

6.6.6 The consumption policy

In this subsection, we simulate a pair sets of gas price paths and oil price paths, then extract the decisions by using the tree algorithm and the LSMC algorithm on the GPU, respectively. Parameters are given in Table 6.1, and we use the market forward curves in Figure 6.2. Figure 6.14 shows the simulated paths and the corresponding consumption policies obtained by using the tree algorithm and the LSMC algorithm, respectively. As we can see, the

FIGURE 6.13: Computing times with respect to mean-reverting rates

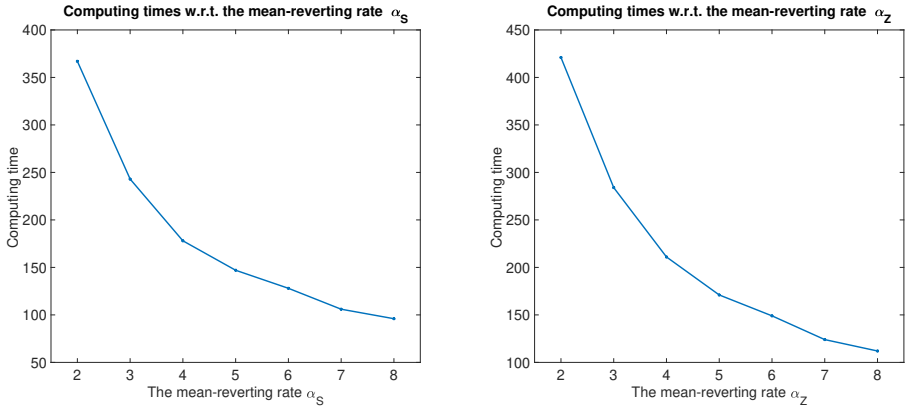


FIGURE 6.14: Paths and decisions

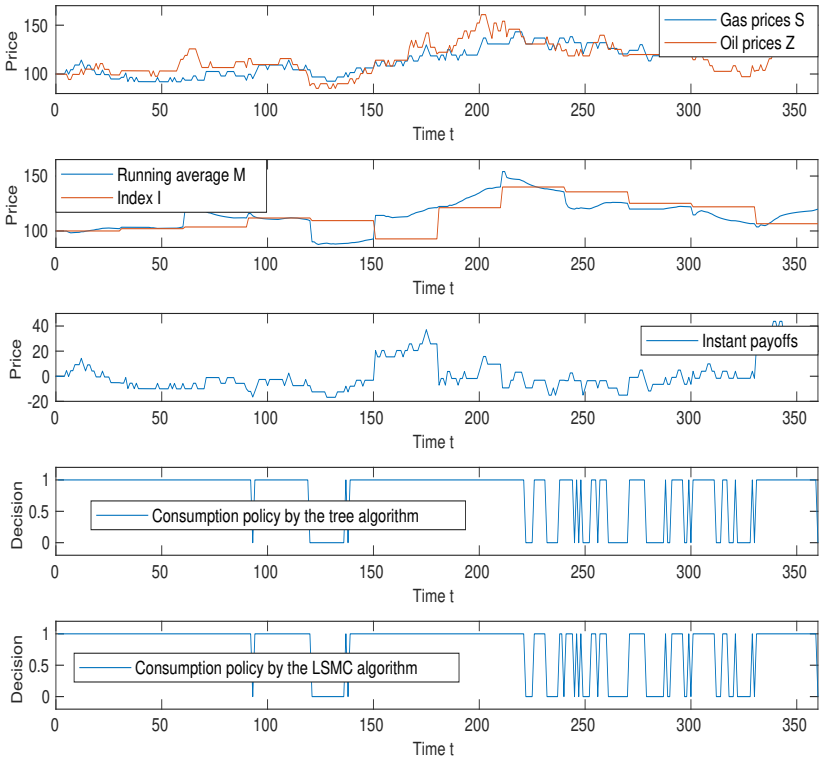
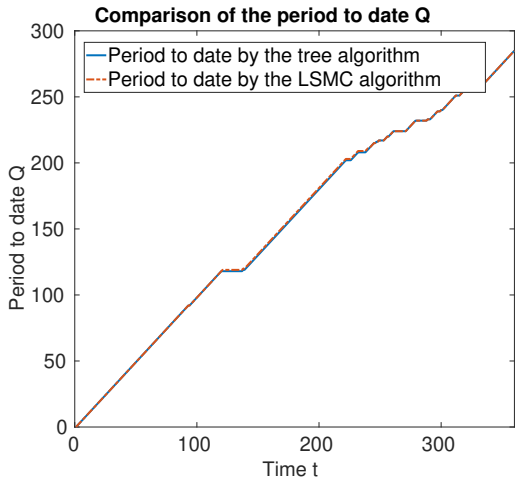


FIGURE 6.15: The period to date Q of different algorithms



Note: In this figure, when using the LSMC algorithm on the GPU, 100000 paths and 1000000 paths have been used in the backward scheme and the forward scheme, respectively. The degree of the polynomials is $h = 3$. When using the tree algorithm, we use the number of points $H = 96$.

tree algorithm and the LSMC algorithm give very similar consumption policies. Generally speaking, both consumption policies suggest taking gas when the instant payoffs are positive and avoiding taking gas when the instant payoffs are negative. However, in the first quarter of the year, both consumption policies suggest taking gas even if the instant payoff is a bit less than zero. This is due to the pressure to avoid penalties. In addition, the total consumptions of both policies are both 285, and both consumption policies suggest to avoid penalties. Figure 6.15 reports the period to date at each time t_n obtained by using these two algorithms. As we can see, the difference between the period to date returned by the tree algorithm and the period to date returned by the LSMC algorithm is quite minimal. Figure 6.15, together with Figure 6.14, show that the tree algorithm and the LSMC algorithm give the same decisions on most days of the year. The tree algorithm suggests taking less gas than the LSMC algorithm on day 120. The difference is compensated for, however, when the LSMC algorithm suggests taking less gas on day 240.

Also, the contract values of the simulated paths obtained by the tree algorithm and the LSMC algorithm are very close: the tree algorithm gives the value 904.91 while the LSMC algorithm gives the value 900.61.

6.7 Conclusion

This chapter has presented numerical examples using all the algorithms built in Chapter 4 and Chapter 5. Numerical examples show that the LSMC algorithm on the GPU is more efficient than the tree algorithm and the LSMC algorithm on the CPU. However, the contract value obtained by LSMC algorithm on the GPU is less optimal due to the loss of information. Although one can use fewer basis functions and simulated paths to ease this loss of information, based on our numerical analysis in Section 5.8, fewer basis functions and simulated paths lead to suboptimal and unstable results. The LSMC algorithm on the CPU does not suffer from this loss of information, but it is quite time consuming, requiring hours to evaluate a one-year contract. The tree algorithm can be implemented efficiently on the GPU with the help of CUDA programming. The contract value obtained by the tree algorithm is sufficiently accurate when compared with the benchmark value. In addition, our numerical examples show that the tree algorithm gives satisfactory values when not too many points are used in the non-uniform grid. The tree algorithm also has the advantage that, for fixed inputs, it gives a fixed contract value. Hence, it is a powerful tool to analysis the contract features of GSAs with indexation. Using the tree algorithm and a number of numerical studies, we demonstrate various features of this complex contract. For example, the contract value and the exercise threshold are more sensitive with respect to the oil price than the value of the running average. The running average and the oil price have less influence on the contract value when their values are large.

Being a lattice-based method, the tree algorithm can be very inefficient when introducing further dimensions. One can easily modify the LSMC algorithm to accommodate basically any sophisticated models, such as multi-factor models, but it is hard to do the same thing for a lattice-based method. Thus, the advantage of the LSMC algorithm is its flexibility. When switching between different models, the only modification that needs to be addressed

is the simulation of the underlying prices. This is why the LSMC algorithm is popular among practitioners. To our knowledge, however, no lattice-based method has been built in the literature that deals with the real GSA contract with indexation. Hence, the tree algorithm provides a trustworthy benchmark for practitioners to test their results.

Chapter 7

Conclusion and future work

In this thesis, we have explored numerical methods for the evaluation of the gas sales agreement where the make-up bank, the carry-forward bank and the indexation are introduced.

The thesis can be divided into two parts. The first part includes Chapter 3, in which we propose a two-dimensional trinomial tree framework for pricing multiple year GSAs with make-up, carry-forward and stochastic strike prices, given the knowledge of forward price dynamics of both gas and index. In the evaluation, we keep track of multiple variables on a daily basis over a number of years and we are able to evaluate efficiently the prices of the contracts so as to find both the optimal daily decisions and the optimal yearly use of both the make-up bank and carry-forward bank. We provide numerical studies in which we compare GSAs with constant strike prices and GSAs with stochastic strike prices. We also demonstrate various features of this complex contract with the help of a number of numerical studies, especially the impact of the make-up and carry-forward banks on both the contract values and decisions.

The second part includes Chapter 4, Chapter 5 and Chapter 6, in which we investigate the evaluation of GSAs with indexation. In Chapter 4, we model the index of a GSA contract by using a running average variable and build a tree algorithm to evaluate GSAs with indexation. In this chapter, the GSA with indexation is also modelled in continuous time and we further prove that the tree-based algorithm can achieve first order consistency to the continuous model. In Chapter 5, we build the least-squares Monte Carlo algorithm for the evaluation of GSAs with indexation and analyze the performance of the LSMC algorithm. Through numerical studies we conclude that

the commonly used regression method regressing on both the gas price and the index underestimates the GSA contract value. We also suggest regressing on the gas price, the oil price, the value of the running average and the index to get better results. In Chapter 6, we introduce basic programming knowledge of general-purpose GPU computing and implement the tree algorithm on the GPU and the LSMC algorithm on both the CPU and GPU. From the numerical experiment, the tree algorithm on the GPU is accurate and efficient while the LSMC algorithm on the CPU is time-consuming. The LSMC on the GPU is the most efficient but suffers from loss of information at runtime. The value surfaces and decision surfaces are presented through the tree algorithm, and various features of the GSA with indexation are discussed in this chapter.

It should be noted that there are other types of indexation in the market. For example, there can be a lag between the month where the index is used and the month where the index is calculated. The index can also be calculated on more than one energy product. These properties make the evaluation of a real contract much harder than the scenario envisaged in this thesis, but we leave this problem to future research. In addition, we do not use parameters estimated from the market data in this thesis. The parameter estimation based on market data itself is a huge topic, which can not be the focus of this thesis but do need more proper and thorough study and analysis, hence we leave it for future research.

Bibliography

Abbas-Turki, L. A. and Lapeyre, B. (2009). "American options pricing on multi-core graphic cards". In: *2009 International Conference on Business Intelligence and Financial Engineering*, Beijing, China, pp. 307–311.

Aleksandrov, N. and Hambly, B. M. (2010). "A dual approach to multiple exercise option problems under constraints". *Mathematical Methods of Operations Research* 71.3, pp. 503–533.

Asche, F., Osmundsen, P. and Tveterås, R. (2002). "European market integration for gas? Volume flexibility and political risk". *Energy Economics* 24.3, pp. 249 –265.

Bardou, O., Bouthemy, S. and Pagès, G. (2009). "Optimal quantization for the pricing of swing options". *Applied Mathematical Finance* 16.2, pp. 183–217.

— (2010). "When are swing options bang-bang?" *International Journal of Theoretical and Applied Finance* 13.6, pp. 867–899.

Barrera-Esteve, C., Bergeret, F., Dossal, C., Gobet, E., Meziou, A., Munos, R. and Reboul-Salze, D. (2006). "Numerical methods for the pricing of swing options: a stochastic control approach". *Methodology and Computing in Applied Probability* 8.4, pp. 517–540.

Basei, M., Cesaroni, A. and Vargiolu, T. (2014). "Optimal exercise of swing contracts in energy markets: an integral constrained stochastic optimal control problem". *SIAM Journal on Financial Mathematics* 5.1, pp. 581–608.

Bender, C. (2011). "Dual pricing of multi-exercise options under volume constraints". *Finance and Stochastics* 15.1, pp. 1–26.

Benguigui, M. and Baude, F. (2014). "Fast American basket option pricing on a multi-GPU cluster". In: *22nd High Performance Computing Symposium*, Tampa, FL, United States, pp. 1–8.

Benth, F. E., Lempa, J. and Nilssen, T. K. (2012). "On the optimal exercise of swing options in electricity markets". *Journal of Energy Markets* 4.4, pp. 3–28.

Bernhart, M. (2011). "Modelization and valuation methods of gas contracts: stochastic control approaches". PhD thesis. Université Paris-Diderot-Paris VII.

Bernhart, M., Tankov, P. and Warin, X. (2011). "A finite-dimensional approximation for pricing moving average options". *SIAM Journal on Financial Mathematics* 2, pp. 989–1013.

Breslin, J., Clewlow, L., Strickland, C. and van der Zee, D. (2008). "Swing contracts: take it or leave it". *Energy Risk* 2, pp. 64–68.

Brigo, D. and Mercurio, F. (2006). *Interest rate models—theory and practice: with smile, inflation and credit*. 2nd Edition. Springer.

Broadie, M. and Cao, M. (2008). "Improved lower and upper bound algorithms for pricing American options by simulation". *Quantitative Finance* 8.8, pp. 845–861.

Broadie, M. and Glasserman, P. (2004). "A stochastic mesh method for pricing high-dimensional American options". *Journal of Computational Finance* 7.4, pp. 35–72.

Chiarella, C., Clewlow, L. and Kang, B. (2016). "The evaluation of multiple year gas sales agreement with regime switching". *International Journal of Theoretical and Applied Finance* 19.1, pp. 1–25.

Clément, E., Lamberton, D. and Protter, P. (2002). "An analysis of a least squares regression method for American option pricing". *Finance and Stochastics* 6.4, pp. 449–471.

Clewlow, L. and Strickland, C. (1999). "Valuing energy options in a one factor model fitted to forward prices". Available online at `https://ssrn.com/abstract=160608`. Last accessed: 14/11/2018.

Clewlow, L., Strickland, C. and Kaminski, V. (2001). "Risk analysis of swing contracts". *Energy and Power Risk Management* 6.5, pp. 32–33.

Curran, M. (2001). "Willow power: optimizing derivative pricing trees". *Algo Research Quarterly* 4.4, pp. 15–23.

Cvetanoska, V. and Stojanovski, T. (2012). "Using high performance computing and Monte Carlo simulation for pricing American options". Available online at `https://arxiv.org/abs/1205.0106`. Last accessed: 14/11/2018.

Dai, M., Li, P. and Zhang, J. E. (2010). "A lattice algorithm for pricing moving average barrier options". *Journal of Economic Dynamics and Control* 34.3, pp. 542 –554.

Dang, D. M., Christara, C. C. and Jackson, K. R. (2009). "A parallel implementation on GPUs of ADI finite difference methods for parabolic PDEs with applications in finance". *Canadian Applied Mathematics Quarterly (CAMQ)* 17.4, pp. 627–660.

— (2012). "An efficient graphics processing unit-based parallel algorithm for pricing multi-asset American options". *Concurrency and Computation: Practice and Experience* 24.8, pp. 849–866.

Dirnstorfer, S., Grau, A. J. and Zagst, R. (2013). "High-dimensional regression on sparse grids applied to pricing moving window Asian options". *Open Journal of Statistics* 3.6, pp. 427–440.

Dong, W. and Kang, B. (2018). "Analysis of a multiple year gas sales agreement with make-up, carry-forward and indexation". *Energy Economics*. Available online at `https://doi.org/10.1016/j.eneco.2018.04.001`. Last accessed: 14/11/2018.

Dörr, U. (2003). "Valuation of swing options and examination of exercise strategies by Monte Carlo techniques". MA thesis. University of Oxford.

Edoli, E. (2013). "Pricing of gas swing contracts with indexed strike: a viscosity solution approach with applications". PhD thesis. University of Padova.

Edoli, E., Fiorenzani, S. and Vargiolu, T. (2016). *Optimization methods for gas and power markets: theory and cases*. Palgrave Macmillan UK.

Edoli, E. and Vargiolu, T. (2013). "Stochastic optimization for the pricing of structured contracts in energy markets". *ARGO Magazine* 2, pp. 35–44.

Edoli, E., Fiorenzani, S., Ravelli, S. and Vargiolu, T. (2013). "Modeling and valuing make-up clauses in gas swing contracts". *Energy Economics* 35, pp. 58 –73.

Egloff, D. (2012). "Pricing financial derivatives with high performance finite difference solvers on GPUs". In: *GPU Computing Gems Jade Edition*. Boston: Morgan Kaufmann, pp. 309 –322.

Fatica, M. and Phillips, E. (2013). "Pricing American options with least squares Monte Carlo on GPUs". In: *Proceedings of the 6th Workshop on High Performance Computational Finance.* Available online at `https://dl.ac`

m.org/citation.cfm?doid=2535557.2535564. Last accessed: 14/11/2018.

Fleming, W. H. and Soner, H. M. (2006). *Controlled Markov processes and viscosity solutions*. 2nd edition. Springer.

Ganesan, N., Chamberlain, R. D. and Buhler, J. (2009). "Acceleration of binomial options pricing via parallelizing along time-axis on a GPU". In: *Proceedings of Symposium on Application Accelerators in High Performance Computing*. Available online at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.395.9455&rep=rep1&type=pdf. Last accessed: 14/11/2018.

Glasserman, P. (2003). *Monte Carlo methods in financial engineering*. Springer.

Gohara, D., Shi, G. and Stone, J. E. (2010). "OpenCL: a parallel programming standard for heterogeneous computing systems". *Computing in Science Engineering* 12.3, pp. 66–73.

Grau, A. J. (2008). "Applications of least-squares regressions to pricing and hedging of financial derivatives". PhD thesis. Technical University of Munich.

Hanfeld, M. and Schlüter, S. (2017). "Operating a swing option on today's gas markets – How least squares Monte Carlo works and why it is beneficial". *Zeitschrift für Energiewirtschaft* 41.2, pp. 137–150.

Hinz, J. (2006). "Valuing virtual production capacities on flow commodities". *Mathematical Methods of Operations Research* 64.2, pp. 187–209.

Holden, L., Løland, A. and Lindqvist, O. (2011). "Valuation of long-term flexible gas contracts". *Journal of Derivatives* 18.3, pp. 75–85.

Hull, J. C. (2011). *Option, futures and other derivatives*. 8th edition. Prentice Hall.

Hull, J. C. and White, A. D. (1993). "Efficient procedures for valuing European and American path-dependent options". *Journal of Derivatives* 1.1, pp. 21–31.

— (1994a). "Numerical procedures for implementing term structure models I: single-factor models". *Journal of Derivatives* 2.1, pp. 7–16.

— (1994b). "Numerical procedures for implementing term structure models II: two-factor models". *Journal of Derivatives* 2.2, pp. 37–48.

Ibáñez, A. (2004). "Valuation by simulation of contingent claims with multiple early exercise opportunities". *Mathematical Finance* 14.2, pp. 223–248.

Ibáñez, A. and Zapatero, F. (2004). "Monte Carlo valuation of American options through computation of the optimal exercise frontier". *Journal of Financial and Quantitative Analysis* 39.2, pp. 253–275.

in't Hout, K. J. and Foulon, S (2010). "ADI finite difference schemes for option pricing in the Heston model with correlation". *International Journal of Numerical Analysis and Modeling* 7.2, pp. 303–320.

Jaillet, P., Ronn, E. I. and Tompaidis, S. (2004). "Valuation of commodity-based swing options". *Management Science* 50.7, pp. 909–921.

Jauvion, G. and Nguyen, T. (2008). *Parallelized trinomial option pricing model on GPU with CUDA*. Available online at `https://www.scribd.com/document/74126721/parallelized-trinomial-option-pricing-model-on-gpu-with-cuda`. Last accessed: 14/11/2018.

Judd, K. L. (1998). *Numerical methods in economics*. MIT press.

Kao, C.-H. and Lyuu, Y.-D. (2003). "Pricing of moving-average-type options with applications". *Journal of Futures Markets* 23.5, pp. 415–440.

Klassen, T. (2001). "Simple, fast and flexible pricing of Asian options". *Journal of Computational Finance* 4.3, pp. 89–124.

Lari-Lavassani, A., Simchi, M. and Ware, A. (2001). "A discrete valuation of swing options". *Canadian Applied Mathematics Quarterly* 9.1, pp. 35–73.

Leitao, Á. and Oosterlee, C. W. (2015). "GPU acceleration of the stochastic grid bundling method for early-exercise options". *International Journal of Computer Mathematics* 92.12, pp. 2433–2454.

Longstaff, F. A. and Schwartz, E. S. (2001). "Valuing American options by simulation: a simple least-squares approach". *Review of Financial Studies* 14.1, pp. 113–147.

Maruyama, G. (1955). "Continuous Markov processes and stochastic equations". *Rendiconti del Circolo Matematico di Palermo* 4.1, pp. 48–90.

Meinshausen, N. and Hambly, B. M. (2004). "Monte Carlo methods for the valuation of multiple-exercise options". *Mathematical Finance* 14.4, pp. 557–583.

Moreno, M. and Navas, J. F. (2003). "On the robustness of least-squares Monte Carlo (LSM) for pricing American derivatives". *Review of Derivatives Research* 6.2, pp. 107–128.

NVIDIA (2018a). "CUBLAS library". Available online at `https://docs.nvidia.com/cuda/cublas/index.html`. Last Accessed: 14/11/2018.

NVIDIA (2018b). "CUDA compiler driver NVCC". Available online at `https://docs.nvidia.com/cuda/cuda-compiler-driver-nvcc/index.html`. Last accessed: 14/11/2018.

— (2018c). "NVIDIA CUDA C programming guide". Available online at `https://docs.nvidia.com/cuda/cuda-c-programming-guide/index.html`. Last Accessed: 14/11/2018.

Øksendal, B. (2003). *Stochastic differential equations: an introduction with applications*. Springer.

Pagès, G. and Wilbertz, B. (2012). "GPGPUs in computational finance: massive parallel computing for American style options". *Concurrency and Computation: Practice and Experience* 24.8, pp. 837–848.

Protter, P. E. (2005). *Stochastic integration and differential equations*. Springer.

Schwartz, E. and Smith, J. E. (2000). "Short-term variations and long-term dynamics in commodity prices". *Management Science* 46.7, pp. 893–911.

Stentoft, L. (2004). "Assessing the least squares Monte-Carlo approach to American option valuation". *Review of Derivatives Research* 7.2, pp. 129–168.

Thanawalla, R. K. (2006). "Valuation of gas swing options using an extended least squares Monte Carlo algorithm". PhD thesis. Heriot-Watt University.

Thompson, A. C. (1995). "Valuation of path-dependent contingent claims with multiple exercise decisions over time: the case of take-or-pay". *Journal of Financial and Quantitative Analysis* 30.2, pp. 271–293.

Wahab, M. I. M. and Lee, C.-G. (2011). "Pricing swing options with regime switching". *Annals of Operations Research* 185.1, pp. 139–160.

Warin, X. (2012). "Hedging swing contract on gas markets". Available online at `https://arxiv.org/abs/1208.5303`. Last accessed: 14/11/2018.

Xu, W., Hong, Z. and Qin, C. (2013). "A new sampling strategy willow tree method with application to path-dependent option pricing". *Quantitative Finance* 13.6, pp. 861–872.