Numerical Methods for Optimal Stochastic Control in Finance

by

Zhuliang Chen

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

Waterloo, Ontario, Canada, 2008

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

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

Abstract

In this thesis, we develop partial differential equation (PDE) based numerical methods to solve certain optimal stochastic control problems in finance.

The value of a stochastic control problem is normally identical to the viscosity solution of a Hamilton-Jacobi-Bellman (HJB) equation or an HJB variational inequality. The HJB equation corresponds to the case when the controls are bounded while the HJB variational inequality corresponds to the unbounded control case. As a result, the solution to the stochastic control problem can be computed by solving the corresponding HJB equation/variational inequality as long as the convergence to the viscosity solution is guaranteed.

We develop a unified numerical scheme based on a semi-Lagrangian timestepping for solving both the bounded and unbounded stochastic control problems as well as the discrete cases where the controls are allowed only at discrete times. Our scheme has the following useful properties: it is unconditionally stable; it can be shown rigorously to converge to the viscosity solution; it can easily handle various stochastic models such as jump diffusion and regime-switching models; it avoids Policy type iterations at each mesh node at each timestep which is required by the standard implicit finite difference methods.

In this thesis, we demonstrate the properties of our scheme by valuing natural gas storage facilities—a bounded stochastic control problem, and pricing variable annuities with guaranteed minimum withdrawal benefits (GMWBs)—an unbounded stochastic control problem. In particular, we use an impulse control formulation for the unbounded stochastic control problem and show that the impulse control formulation is more general than the singular control formulation previously used to price GMWB contracts.

Acknowledgements

First of all, I would like to dedicate my wholehearted thanks to my supervisor Prof. Peter Forsyth. His flawless guidance "safely" led me through three-year's academia life. His infinite enthusiasm on research inspired me all the time. His everyday discussions and chats exposed me to all kinds of interesting problems, research methodologies and principles. His smart ideas always surprised me. His more-than-sufficient financial supports eliminated common worrisome real-life issues for me. I am really lucky to be able to study with Prof. Forsyth. My Ph.D. life was a very happy and satisfactory journey.

I would like to greatly thank my Ph.D. committee members: Prof. George Labahn, Prof. Yuying Li, Prof. Andrew Heunis as well as my external examiner Prof. Tony Ware. Thank you very much for carefully reading my thesis and providing useful comments and advices. I also appreciate a lot of helpful discussions with Prof. Yuying Li on my research problems.

I would like to thank Prof. Arne Storjohann, Prof. Justin Wan, Prof. Jeff Orchard for many discussions and interesting chats, especially those on babies.

My wife Kunling Weng worths my countless thanks. My Ph.D. work was undoubtedly due to the love and care with all her heart. She also took care our little baby herself without a single complaint.

I would like to greatly thank my parents for their best support, helpful advices and absolute trust.

I would like to give plenty of appreciations to all my friends in Waterloo, especially Chen Yu, Liu Yinbin, Zhou Wei, Wang Jian, Wan Weihan, Li Wei, Chen Jun, Zhang Jie for their warmhearted helps. Many thanks to all my SciCom members especially Mario, Ruonan, Shannon, Simon, Amelie, Dashan, Lin, Iris for their helps and interesting chats.

Dedication

To Runxi. Happy Birthday!

Contents

List of Tables

List of Figures

List of Algorithms

Chapter 1

Introduction

Many problems in finance are characterized as optimal stochastic control problems. Examples discussed in this thesis include the valuation of natural gas storage facilities and the valuation of variable annuities with guaranteed minimum withdrawal benefits (GMWBs). Some other examples of stochastic control in finance include portfolio selection problems with liquidity risk and price impact [2, 83], with transaction costs [32, 70, 59], or with capital gains taxes [79]; optimal hedging with transaction costs [33, 62, 68]; financial hedging of operational risk [66]; and valuation of power generation assets [81, 24]. Refer to [72] for a survey of various stochastic control problems and the applications in finance.

A typical stochastic control problem in finance can be thought of as consisting of two components: an underlying stochastic model (e.g., a mean-reverting process for natural gas prices) and a model of an inventory variable (e.g., amount of gas in storage). The two components are affected by different control strategies (e.g., withdrawing gas from the storage reduces the gas inventory in storage). At each point in state space, an optimal control needs to be determined to maximize some objective function (e.g., expected revenues that the operators of the gas storage facility can obtain by optimally operating the facility).

Numerical methods which have been developed to solve stochastic control problems include Markov chain based methods (see, e.g., [61, 17]), simulation based methods (see, e.g., [19, 12, 82]), lattice based approaches (see, e.g., [38]), and partial differential equation

(PDE) based approaches (see, e.g., [22, 23, 80, 46, 64, 31]).

Markov chain, simulation and lattice based approaches solve the stochastic control problems using dynamic programming. Lattice based approaches are essentially explicit finite difference methods, which thus suffer from timestep restrictions due to stability conditions. Simulation based methods are well suited for solving multi-dimensional problems (e.g., problems with five state variables) which the PDE based approaches cannot handle. Nevertheless, it is known that such methods have difficulty achieving high accuracy. Furthermore, if the optimal controls are not a finite set, then simulation based methods will have to approximate the controls as piece-wise constant, which will substantially increase the computational cost.

As shown in [72, 45], the value of a stochastic control problem is normally identical to the viscosity solution of a Hamilton-Jacobi-Bellman (HJB) equation or an HJB variational inequality. The HJB equation corresponds to the case when the controls are bounded while the HJB variational inequality corresponds to the unbounded control case. As a result, the solution to the stochastic control problem can be computed by solving the corresponding HJB equation/variational inequality using PDE based approaches. In general, the solution to the HJB equation/variational inequality may not be unique. As noted in [46, 6], it is important to ensure that a numerical scheme converges to the viscosity solution of the equation, which is normally the appropriate solution of the corresponding stochastic control problem.

Provided a strong comparison result for the HJB equation/variational inequality applies, the authors of [11, 6] demonstrate that a numerical scheme will converge to the viscosity solution of the equation if it is l_∞-stable, consistent, and monotone. Schemes failing to satisfy these conditions may converge to non-viscosity solutions. In fact, the authors of [75] give an example where seemingly reasonable discretizations of nonlinear option pricing PDEs that do not satisfy the sufficient convergence conditions for viscosity solutions either never converge or converge to a non-viscosity solution.

Consequently, our research focuses on developing PDE-based numerical methods with the following desired properties for solving optimal stochastic control problems:

1. No timestep restrictions should be imposed.

2. The methods should converge to the viscosity solution.

3. The methods should be independent of the underlying stochastic models so that they can be easily applied for different stochastic models such as mean-reverting models, geometric Brownian motion (GBM) models, regime-switching models and jump diffusion models.

4. The methods should be at least as efficient as other existing methods.

1.1 Contributions

We first consider the stochastic control problem with bounded controls, which corresponds to an HJB equation. The HJB equation is normally convection dominated, that is, the equation has no diffusion in the inventory component direction. Hence it is well known to be difficult to solve numerically.

As such, we develop a fully implicit scheme based on a semi-Lagrangian timestepping to discretize the HJB equation. Initially introduced by [42, 74] for atmospheric and weather numerical predictions, semi-Lagrangian schemes can effectually reduce the numerical problems arising for convection dominated equations. The fully implicit semi-Lagrangian scheme meets our goal in the sense that it satisfies each of the above properties.

We demonstrate the properties of our scheme by valuing gas storage cash flows in Chapters 2- 6. Our work in this area makes the following contributions:

- We develop fully implicit and Crank-Nicolson finite difference schemes based on a semi-Lagrangian method for solving the HJB equation for the gas storage problem, assuming the gas spot price follows a one-factor mean-reverting model, and obtain the optimal control strategies (see Chapter 2).

- We show that the fully implicit, semi-Lagrangian discretization is algebraically identical to a discretization based on a scenario that the operations on the storage are performed only at discrete times (see Chapter 2).

- We show that, compared to the standard implicit methods given in [46], semi-Lagrangian timestepping methods avoid the need for Policy type iterations at each node at each timestep. Instead, the methods require solution of a discrete local optimization problem at each mesh point in order to determine the optimal control value. The optimization problem can be solved efficiently so that the complexity per timestep of the fully implicit semi-Lagrangian scheme is the same as that of an explicit method or the complexity per iteration per timestep of a standard implicit method (see Chapter 2).

- We give an intuitive introduction to the notation of (possibly discontinuous) viscosity solutions that is able to handle various types of boundary conditions (see Chapter 3).

- We prove that the fully implicit semi-Lagrangian scheme is unconditionally l_∞-stable, monotone and consistent. Therefore, provided a strong comparison property holds, the fully implicit, semi-Lagrangian discretization converges to the unique and continuous viscosity solution of the gas storage pricing equation using the results in [11, 6] (see Chapter 3).

- Numerical experiments indicate that fully implicit timestepping can achieve first-order convergence, while Crank-Nicolson timestepping does not appear to converge at a higher than first-order rate. Thus fully implicit timestepping is probably a better choice since it guarantees convergence to the viscosity solution and it is also straightforward to implement (see Chapter 4).

- We propose a one-factor regime-switching model for the risk neutral natural gas spot price. We calibrate model parameters to both market futures and options. The calibration results suggest that the regime-switching model is capable of fitting

the market data more accurately than the one-factor mean-reverting model (see Chapter 5).

- Since the semi-Lagrangian timestepping completely separates the inventory component from the underlying stochastic model, we can extend our scheme for gas storage problem under a one-factor mean-reverting model to solve the problem under the regime-switching model, using the model parameters obtained from the calibration. Provided a unique continuous viscosity solution exists, we prove the convergence of the scheme to the viscosity solution using the results in [71, 11, 6] (see Chapter 6).

We then consider the stochastic control problem with unbounded controls. We form an impulse control formulation for the unbounded control problem, resulting in an HJB variational inequality. We then solve the HJB variational inequality using an extension of the semi-Lagrangian method for the bounded control problems.

We demonstrate the properties of the semi-Lagrangian timestepping for the unbounded control case by pricing the value of GMWB variable annuities in Chapters 7- 9. Our main contributions in this area are summarized as follows:

- We first study the GMWB variable annuity pricing problem under the scenario that the controls (i.e., withdrawing funds) are allowed only at discrete times. We formulate a pricing model for the problem. We then present a numerical scheme for solving the pricing model and prove that the scheme converges to the unique viscosity solution of the problem (see Chapter 7).

- We then consider the GMWB valuation problem assuming the operations are allowed continuously. We propose an impulse control formulation for this problem, resulting in an HJB variational inequality (see Chapter 8).

- We show that the impulse control formulation is a more general approach compared to the singular control formulation previously used to price the GMWB contracts (see Chapter 8).

- We generalize the scheme for the discrete control case to handle the continuous control case. We show that the scheme is also identical to an extension of the semi-Lagrangian timestepping method for the bounded stochastic control problems (see Chapter 8).

- Provided a strong comparison result holds, we prove that the scheme converges to the unique viscosity solution of the HJB variational inequality corresponding to the impulse control problem by verifying the l_∞ stability, monotonicity and consistency of the scheme and using the results in [11, 6] (see Chapter 8).

- Through numerical experiments, we demonstrate that the solution of our impulse control formulation are the same (to within discretization errors) as that of the singular control formulation. The numerical results also appear to show that the optimal control strategy may not be unique. That is, there exists a region where different control strategies can result in the same contract value (see Chapter 8).

- Using the semi-Lagrangian discretization, we carry out an extensive analysis of the no-arbitrage fee for GMWB guarantees assuming various parameters and contract details. In particular, we consider the effects of incorporating a separate mutual fund management fee, assuming sub-optimal investor behaviours and various contract parameters such as reset provisions, maturities, withdrawal intervals, and surrender charges. Our numerical experiments show that the GMWB insurance fees currently charged by the insurance companies are not enough to cover the costs of hedging these contracts (see Chapter 9).

To summarize, our main result in this thesis is that we have a unified numerical scheme based on a semi-Lagrangian approach that is able to solve both bounded and unbounded stochastic control problems as well as the discrete control cases where the operations are allowed only at discrete times. In addition, our numerical scheme is shown to converge to the viscosity solution of these problems.

1.2 Outline

The rest of this thesis is arranged as follows. In Chapter 2 we propose the semi-Lagrangian schemes for valuing gas storage cash flows. Convergence analysis of the fully implicit, semi-Lagrangian scheme is given in Chapter 3. We conduct numerical experiments for gas storage problem in Chapter 4. Chapter 5 compares a one-factor regime-switching model with a one-factor mean-reverting model through empirical calibration. In Chapter 6 we solve the gas storage problem using the calibrated regime-switching model. Chapter 7 proposes a numerical scheme for pricing GMWB variable annuities assuming withdrawals occur only at predetermined discrete times. In Chapter 8, we generalize the scheme to solve the impulse control problem, corresponding to the GMWB valuation problem under the continuous withdrawal scenario. In Chapter 9 we study the effects of various parameters and contract details on the no-arbitrage fees of GMWB contracts. Finally, conclusions are drawn in Chapter 10.

Chapter 2

Valuation of Natural Gas Storage Facilities and Optimal Operations

In this chapter we develop numerical schemes for pricing the value of natural gas storage facilities, which is characterized as an optimal stochastic control problem with a bounded control. We begin by defining the stochastic control problem, and then reformulate the problem under the partial differential equation (PDE) framework. We then present the numerical schemes based on the semi-Lagrangian method for the pricing PDE.

2.1 Introduction and Previous Work

Similar to other commodities such as fuel and electricity, natural gas prices exhibit seasonality dynamics due to fluctuations in demand [73]. As such, natural gas storage facilities are constructed to provide a cushion for such fluctuations by releasing natural gas in storage in seasons with high demand.

Recently, several authors [1, 80, 84, 85, 64, 18, 12] have discussed the no-arbitrage value of natural gas storage facilities (or, equivalently, the values of gas storage contracts for leasing the storage facilities). The value of a gas storage facility can be regarded as the maximum expected revenues under the risk neutral measure that the operator of the

facility can obtain by optimally operating the facility, that is, "buying low" and "selling high". As a result, the valuation of gas storage facilities is characterized as a stochastic control problem.

Assuming that the control is of bang-bang type (that is, the control takes values only from a finite set), in [18, 12], simulation based methods are used to directly solve the stochastic control problem for gas storage valuation. These methods are well suited for solving multi-dimensional problems (e.g., problems with five state variables) which the PDE-based approaches cannot handle. Nevertheless, it is known that such methods have difficulty achieving high accuracy. Furthermore, if the control is not of bang-bang type, such methods will have to approximate the control as piece-wise constant, which will substantially increase the computational cost. See [24] for descriptions of control problems which are not of the bang-bang type for valuation of electricity power plants.

In [80], an explicit finite difference scheme is used to solve the pricing PDE derived from the stochastic control problem for gas storage valuation. As is well known, explicit timestepping suffers from timestep restrictions due to stability considerations. Alternatively, the authors of [46] present implicit finite difference schemes, which eliminate the timestep restriction, for solving general controlled PDE. However, this scheme requires solution of nonlinear discretized algebraic equations using a Policy type iteration at each timestep. Reference [46] introduces another type of implicit scheme that approximates the control as piece-wise constant to avoid the need for solving nonlinear equations at the expense of solving a number of linear problems at each timestep. Similar to the simulation based methods, if the control is not of bang-bang type, such schemes will be computationally expensive. In [84], a finite element semi-Lagrangian scheme is developed to solve a PDE for certain gas storage contracts. In [85, 64], a wavelet method coupled with a semi-Lagrangian approach is used to solve the gas storage PDE. While the wavelet method shows promise, it is difficult to obtain a rigorous proof of convergence to the viscosity solution.

2.2 The Mathematical Model

In this section we present the mathematical model for valuing natural gas storage facilities. First, we formulate a stochastic control problem for the value of gas storage facilities. Then we assume a one-factor model for natural gas spot prices and a model for the gas inventory. Finally, we heuristically derive the Hamilton-Jacobi-Bellman (HJB) equation from the stochastic control framework using dynamic programming (i.e., Bellman's Principle) and Itô's Lemma.

2.2.1 Problem Notation

We use the following notation for the natural gas storage problem:

- P: the current spot price per unit of natural gas.
- I: current amount of working gas inventory. We assume that I can be any value lying within the domain $[0, I_{\max}]$.
- $\hat{V}(P, I, t)$: value of the natural gas storage facility (e.g., the leasing rate of the facility) with respect to natural gas price P and inventory level I at time t.
- T: expiry time of the contract.
- c: control variable that represents the rate of producing gas from or injecting gas into the gas storage ($c > 0$ represents production, $c < 0$ represents injection). If $c = 0$ then no operation is performed on the storage.
- $c_{\max}(I)$: the maximum rate at which gas can be released from storage as a function of inventory levels, $c_{\max}(I) > 0$. We use the expression in [80]

$$c_{\max}(I) = k_1\sqrt{I}, \tag{2.2.1}$$

  where k_1 is a positive constant. This implies $c_{\max}(0) = 0$ with the physical meaning that no gas can be produced if the gas storage is empty.

- $c_{\min}(I)$: $|c_{\min}(I)|$ is the maximum rate at which gas can be injected into storage as a function of inventory levels. Note that $c_{\min}(I) < 0$, with our sign convention that $c > 0$ represents production. We use the expression from [80]

$$c_{\min}(I) = -k_2\sqrt{\frac{1}{I+k_3} - \frac{1}{k_4}}, \tag{2.2.2}$$

where k_2, k_3 and k_4 are positive constants, and k_2, k_3, k_4 satisfy the constraint $c_{\min}(I_{\max}) = 0$, which means that no gas injection is possible if the gas storage is full. Equation (2.2.2) implies that

$$|c_{\min}(I)| \;\leq\; \left|const.\sqrt{I_{\max} - I}\right| \;\;;\;\; I < I_{\max} \;\;,\;\; I \to I_{\max} \;\;. \tag{2.2.3}$$

- $a(c)$: the rate of gas loss incurred inside the storage given a gas production/injection rate of c. In general, the change in gas inventory satisfies

$$\frac{dI}{dt} = -(c + a(c)), \tag{2.2.4}$$

where usually $a(c) \geq 0$. We use the model in [80]

$$a(c) = \begin{cases} 0 & \text{if } c \geq 0, \text{ (producing gas)}, \\ k_5 & \text{if } c < 0, \text{ (injecting gas)}, \end{cases} \tag{2.2.5}$$

where $k_5 > 0$ is a constant.

- $b(c)$: the rate of gas loss incurred outside of the storage (e.g., gas loss incurred during the transportation process) given a gas production/injection rate of c.

In the rest of the thesis, we will follow [80] and assume $a(c) = b(c)$. Consequently, we will use $a(c)$ to represent the rate of gas loss incurred both inside and outside of the storage. Nevertheless, our theoretical results in this thesis still easily follow for the case when $a(c) \neq b(c)$.

Note that if we are using a control c satisfying $-k_5 < c < 0$, that is, injecting at a rate smaller than the rate of gas loss, then equation (2.2.5) implies that $c + a(c) > 0$. According to equation (2.2.4), this means that injecting natural gas into the gas storage decreases the gas inventory, which is unreasonable. Consequently, we further require that control c satisfies the constraint $c \in [c_{\min}(I), -k_5] \cup [0, c_{\max}(I)]$ so that

$$c + a(c) \leq 0 \quad \text{if } c < 0 \text{ (injecting gas).} \tag{2.2.6}$$

In other words, the operator of the gas storage facility is not allowed to inject and at the same time decrease the gas inventory. We point out that the constraint on the control also makes the boundary conditions well defined (this will be discussed in more detail in a subsequent section). For future reference, given any $I \in [0, I_{\max}]$, we define the set $C(I)$ as

$$C(I) = [c_{\min}(I), -k_5] \cup [0, c_{\max}(I)], \tag{2.2.7}$$

where we adopt the convention that $[\alpha, \beta] = \emptyset$ if $\alpha > \beta$.

2.2.2 Stochastic Control Formulation

Under the stochastic control framework, the value of a gas storage facility at a point (P, I, t) is the maximum expected revenues under risk neutral measure during the period that the storage facility can generate before the contract expires. Therefore, we can write $\hat{V}(P, I, t)$ as

$$\begin{aligned} &\hat{V}(P, I, t) \\ &= \sup_{c(s) \in C(I(s))} E^{\mathbb{Q}} \left[\int_t^T e^{-r(s-t)} \big[c(s) - a(c(s)) \big] P(s) ds + e^{-r(T-t)} \hat{V}\big(P(T), I(T), T\big) \right], \end{aligned} \tag{2.2.8}$$

where

- $P(s)$ is a stochastic gas price path in the time direction with $P(t) = P$.

- $I(s)$ is a stochastic gas inventory path in the time direction with $I(t) = I$.
- $c(s)$ is a control path in the time direction.
- $C(I)$ is given in (2.2.7).
- $E^{\mathbb{Q}}$ is the conditional expectation under risk neutral $\mathbb{Q}$ measure with initial values $P(t) = P$ and $I(t) = I$.
- $r > 0$ is the continuously compounded risk-free interest rate.
- $[c(s) - a(c(s))] \cdot P(s)$ represents the instantaneous rate of revenue obtained at time s by producing natural gas from $(c(s) > 0)$ or injecting gas into $(c(s) < 0)$ the gas storage facility, taking into account the possibility of gas loss outside of the storage (recall that we have assumed $b(c) = a(c)$).
- the integral term represents the total amount of discounted cash flows received during the period $[t, T]$.
- $\hat{V}\big(P(T), I(T), T\big)$ is the contract payoff.

Note that the expectation is taken under risk neutral $\mathbb{Q}$ measure since hedging a gas storage facility is possible [84, 85].

2.2.3 Natural Gas Spot Price Model

In this subsection, we present a one-factor mean-reverting process for natural gas spot price. This process is able to capture the mean-reverting and seasonality effects of natural gas spot prices. However, there is a certain amount of controversy surrounding the precise form of a reasonable natural gas spot price model [73]. In a later chapter, we will develop a regime-switching model for natural gas spot prices which, as demonstrated by empirical calibration, is better than the one-factor mean-reverting model. The numerical methods that we derive in this chapter can easily generalize to more complex spot price models, including the regime-switching model.

Since the expectation in (2.2.8) is taken under risk neutral $\mathbb{Q}$ measure, we directly assume a risk neutral price process for natural gas spot, given by the following stochastic differential equation (SDE):

$$dP = \alpha(K(t) - P)dt + \hat{\sigma}(P)PdZ \tag{2.2.9}$$

$$K(t) = K_0 + \beta_{SA}\sin(4\pi(t - t_{SA})), \tag{2.2.10}$$

where

- $\alpha > 0$ is the mean-reverting rate,
- $K(t) \geq 0$ is the long-term equilibrium price that incorporates seasonality,
- $\hat{\sigma}(P)$ is the volatility as a function of P. We will give more details of this function in Section 2.2.5.
- dZ is an increment of the standard Gauss-Wiener process,
- $K_0 \geq 0$ is the equilibrium price without seasonality effect,
- β_{SA} is the semiannual seasonality parameter,
- t_{SA} is the seasonality centering parameters, representing the time of semiannual peak of equilibrium price in summer and winter.

According to equation (2.2.10), the equilibrium price $K(t)$ is a periodic function with period 1/2. This models the seasonal evolution of the annual equilibrium price, e.g., $K(t)$ exhibits two peaks annually, respectively corresponding to high natural gas spot prices in summer and winter.

2.2.4 Pricing Equation

After writing the stochastic control formulation (2.2.8), the next step is to convert the formulation to a PDE using dynamic programming and Itô's Lemma. Note that the

transform requires that the solution $\hat{V}$ be smooth, which may not be true. However, the value of the stochastic control problem normally coincides with the viscosity solution of the corresponding PDE. For example, the authors of [72, 45, 70, 16, 83, 32, 34, 48] have proved that, for many types of stochastic control problems including both bounded and unbounded controls, the value function is the viscosity solution of the corresponding nonlinear PDE. Proving such an equivalence is, nevertheless, beyond the scope of this thesis. As a result, we will assume in this thesis that the solution of a stochastic control problem is identical to the viscosity solution of the PDE resulting from applying the dynamic programming and Ito's Lemma; we will instead focus on developing numerical schemes for solving the PDE which will converge to the viscosity solution.

A non-rigorous derivation for the PDE is given in Appendix A, which results in the following HJB equation from the stochastic control equation (2.2.8)

$$\hat{V}_t+\frac{1}{2}(\hat{\sigma}(P)P)^2\hat{V}_{PP}+\alpha(K(t)-P)\hat{V}_P+\sup_{c\in C(I)}\left[(c-a(c))P-(c+a(c))\hat{V}_I\right]-r\hat{V}=0. \quad (2.2.11)$$

Note that it is also possible to directly derive (2.2.11) based on a hedging argument [84, 85].

For a financial contract such as the natural gas storage contract, a terminal payoff is given at the maturity $t = T$. In order to compute the value of the contract today, we need to solve the pricing PDE backwards in time from $t = T$ to $t = 0$. Let $\tau = T - t$ denote the current time to maturity. For ease of exposition, we will write our PDE in terms of τ so that we will solve the pricing PDE from $\tau = 0$ to $\tau = T$. Let $V(P, I, \tau)$ denote the value of a natural gas storage facility as a function of (P, I, τ). In terms of the facility value $\hat{V}(P, I, t)$ at forward times with respect to t, we have the identity $V(P, I, \tau) = \hat{V}(P, I, T - \tau) = \hat{V}(P, I, t)$.

Rewriting equation (2.2.11) using the variable τ results in

$$V_\tau=\frac{1}{2}(\hat{\sigma}(P)P)^2V_{PP}+\alpha(K(t)-P)V_P+\sup_{c\in C(I)}\left[(c-a(c))P-(c+a(c))V_I\right]-rV\ , \quad (2.2.12)$$

2.2.5 Boundary Conditions

In order to completely specify the gas storage problem, we need to provide boundary conditions.

A number of terminal boundary conditions can be used. Note that since we will be solving backwards in time, the terminal state occurs at $\tau = 0$. Typical examples include

- A zero payoff, as suggested by [80]: $V(P, I, 0) = 0$.
- A non-negative payoff obtained by selling all the leftover gas in the storage at the maximum rate, that is, $V(P, I, 0)$ is the solution $\bar{V}$ to the PDE obtained by fixing control $c = c_{\max}(I)$ in PDE (2.2.12) and solving the resulting PDE backwards from $\tau = 0$ to $\tau = \infty$ with $\bar{V}(P, I, 0) = 0$. We then specify $V(P, I, 0) = \bar{V}(P, I, \infty)$.
- The penalty payoff introduced by [18]:

$$V(P, I, 0) = const. \cdot P \cdot \min(I - I_0, 0) \ , \tag{2.2.13}$$

  where $const. > 0$ and I_0 represents the inventory level at time $t = 0$. This has the meaning that a penalty will be charged if the gas inventory in storage when the facility is returned is less than the gas inventory at the inception of the contract.

The domain for the PDE (2.2.12) is $P \times I \in [0, \infty] \times [0, I_{\max}]$. For computational purposes, we need to solve the PDE in a finite computational domain $[0, P_{\max}] \times [0, I_{\max}]$.

As $I \to 0$, from equations (2.2.1-2.2.7) we have that

$$c + a(c) \leq 0 \ \ ; \ \ \forall c \in C(I) \ , \ I \to 0 \ . \tag{2.2.14}$$

Hence the characteristics are outgoing (or zero) in the I direction at $I = 0$, and we simply solve the PDE along the $I = 0$ boundary, no further information is needed. Condition (2.2.14) has the interpretation that gas cannot be produced from a facility which is empty.

Similarly, as $I \to I_{\max}$, equations (2.2.1-2.2.7) imply that

$$c + a(c) \geq 0 \quad ; \quad \forall c \in C(I) \ , \ I \to I_{\max}, \tag{2.2.15}$$

which again means that the characteristics are outgoing (or zero) in the I direction at $I = I_{\max}$. Consequently, we simply solve the PDE along the $I = I_{\max}$ boundary, no further information is needed. Condition (2.2.15) has the interpretation that gas cannot be injected into the storage facility when it reaches full capacity.

Taking the limit of equation (2.2.12) as $P \to 0$, we obtain

$$V_\tau \quad = \quad \alpha K(t) V_P + \sup_{c \in C(I)} \left[-(c + a(c)) V_I \right] - rV \quad ; \quad P \to 0 \ . \tag{2.2.16}$$

Since $\alpha K(t) \geq 0$, we can solve (2.2.16) without requiring additional boundary conditions, as we do not need information from outside the computational domain $[0, P_{\max}]$.

In this chapter, we assume that $\hat{\sigma}(P)$ is a continuous function that satisfies $\hat{\sigma}(P) = \sigma$ for $P \in [0, P_{\max} - \epsilon]$, where $\sigma > 0$ is a constant and $\epsilon > 0$ is a constant close to zero. For $P \in [P_{\max} - \epsilon, P_{\max}]$, $\hat{\sigma}(P)$ approaches 0 continuously as $P \to P_{\max}$. In other words, the volatility is constant for most values of P and decreases to zero as $P \to P_{\max}$. The decreasing behaviour of $\hat{\sigma}(P)$ at the far boundary $P = P_{\max}$ generates a negligible error by choosing $P_{\max}$ sufficiently large (see Chapter 4 for numerical results for different choices of $P_{\max}$).

Based on this form of $\hat{\sigma}(P)$, taking the limit of equation (2.2.12) as $P \to P_{\max}$, we have

$$V_\tau = \alpha(K(t) - P) V_P + \sup_{c \in C(I)} \left[(c - a(c)) P - (c + a(c)) V_I \right] - rV \quad ; \quad P \to P_{\max}. \tag{2.2.17}$$

We will always choose $P_{\max} \gg K(t)$, hence equation (2.2.17) can be solved at $P = P_{\max}$ without additional information.

The purpose of introducing the continuous function $\hat{\sigma}(P)$ is so that the boundary equation (2.2.17) is the limit of the PDE from the domain interior. This will reduce the

technical manipulations required to prove convergence of our numerical scheme to the viscosity solution of the modified gas storage problem (2.2.12–2.2.17) (see Chapter 3). Since $\epsilon \ll 1$, the numerical implementation assuming that $\hat{\sigma}(P)$ has the above behaviour is, for all practical purposes, the same as an implementation assuming that $\hat{\sigma}(P) = \sigma$ for $P < P_{\max}$ and $\hat{\sigma}(P) = 0$ for $P = P_{\max}$. This has the intuitive interpretation of specifying the commonly used boundary condition $V_{PP} = 0$, $P \to \infty$.

2.3 Numerical Methods Based on Semi-Lagrangian Timestepping

A semi-Lagrangian approach is presented in [40] for pricing continuously observed American Asian options under jump diffusion. In this section, we extend the semi-Lagrangian method in [40] to solve the HJB equation (2.2.12) and associated boundary conditions (2.2.13-2.2.17) that involve optimal control. The main idea used to construct a semi-Lagrangian discretization of the PDE (2.2.12) is to integrate the PDE along a semi-Lagrangian trajectory (defined below). Various semi-Lagrangian discretizations can be obtained by evaluating the resulting integrals using different numerical integration methods: for example, using the rectangular rule provides a fully implicit timestepping scheme, while using the trapezoidal rule gives a Crank-Nicolson timestepping scheme.

This section is arranged as follows: we first present an intuitive idea for developing a semi-Lagrangian discretization for equation (2.2.12). We then present both a fully implicit and a Crank-Nicolson timestepping scheme based on this idea.

We will show that the fully implicit semi-Lagrangian scheme is identical to a scheme derived based on a purely physical reasoning, described in Appendix C, under the scenario that the operator of a gas storage facility can change the controls only at fixed discrete times. This ensures that the fully implicit semi-Lagrangian scheme satisfies discrete no-arbitrage jump conditions. The correspondence between the fully implicit semi-Lagrangian discretization and the discrete control problem also holds for other applications.

At the end of this section, we reformulate the discrete equations into a matrix form and present an algorithm to compute the solution.

Prior to presenting the timestepping schemes, we introduce the following notation. We use an unequally spaced grid in P direction for the PDE discretization, represented by $[P_0, P_1, \ldots, P_{i_{\max}}]$. Similarly, we use an unequally spaced grid in the I direction denoted by $[I_0, I_1, \ldots, I_{j_{\max}}]$. We denote by $0 < \Delta\tau <, \ldots, < N\Delta\tau = T$ the discrete timesteps used to discretize the PDE (2.2.12). Let $\tau^n = n\Delta\tau$ denote the nth timestep. It will be convenient to define $\Delta P_{\max} = \max_i\big(P_{i+1} - P_i\big)$, $\Delta P_{\min} = \min_i\big(P_{i+1} - P_i\big)$, $\Delta I_{\max} = \max_j\big(I_{j+1} - I_j\big)$, $\Delta I_{\min} = \min_j\big(I_{j+1} - I_j\big)$. We assume that there are mesh size/timestep parameters h such that

$$\Delta P_{\max} = C_1 h \;\; ; \;\; \Delta I_{\max} = C_2 h \;\; ; \;\; \Delta\tau = C_3 h \;\; ; \Delta P_{\min} = C_1' h \;\; ; \;\; \Delta I_{\min} = C_2' h. \quad (2.3.1)$$

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

Let $V(P_i, I_j, \tau^n)$ denote the exact solution of equation (2.2.12) when the natural gas spot price resides at node P_i, the gas inventory at node I_j and discrete time at τ^n. Let $V_{i,j}^n$ denote an approximation of the exact solution $V(P_i, I_j, \tau^n)$. Let $\mathcal{L}$ be a differential operator represented by

$$\mathcal{L}V = \frac{1}{2}\hat{\sigma}^2(P)P^2 V_{PP} + \alpha(K(t) - P)V_P - rV. \quad (2.3.2)$$

Using the differential operator (2.3.2), we can rewrite the natural gas storage pricing PDE (2.2.12) as

$$\inf_{c \in C(I)} \{V_\tau + (c + a(c))V_I - (c - a(c))P - \mathcal{L}V\} = 0. \quad (2.3.3)$$

We use standard finite difference methods to discretize the operator $\mathcal{L}V$. Let $(\mathcal{L}_h V)_{i,j}^n$ denote the discrete value of the differential operator (2.3.2) at node (P_i, I_j, τ^n). The operator (2.3.2) can be discretized using central, forward, or backward differencing in the P direction to give

$$(\mathcal{L}_h V)_{i,j}^n = \gamma_i^n V_{i-1,j}^n + \beta_i^n V_{i+1,j}^n - (\gamma_i^n + \beta_i^n + r)V_{i,j}^n, \quad (2.3.4)$$

where γ_i^n and β_i^n are determined using Algorithm B.1 given in Appendix B. The algorithm guarantees γ_i^n and β_i^n satisfy the following positive coefficient condition:

$$\gamma_i^n \geq 0 \quad ; \quad \beta_i^n \geq 0 \quad i = 0, \ldots, i_{\max} \; ; \; j = 0, \ldots, j_{\max} \; ; \; n = 1, \ldots, N. \tag{2.3.5}$$

As we will demonstrate in Section 3.2, the positive coefficient property (2.3.5) is sufficient to ensure convergence of a semi-Lagrangian fully implicit timestepping scheme to the viscosity solution of the HJB equation (2.2.12). All our discretizations presented in this thesis are assumed to satisfy the positive coefficient condition.

2.3.1 An Intuitive Derivation

Now we give the intuition for developing the semi-Lagrangian discretization schemes. Let us consider a path (or a semi-Lagrangian trajectory) $\mathcal{I}(\tau)$ that follows the ODE

$$\frac{d\mathcal{I}}{d\tau} = c + a(c). \tag{2.3.6}$$

According to (2.3.6), we can write the term $V_\tau + (c + a(c))V_I$ in (2.3.3) in the form of a Lagrangian directional derivative

$$\frac{DV}{D\tau} = \frac{\partial V}{\partial \tau} + \frac{\partial V}{\partial I}\frac{d\mathcal{I}}{d\tau}. \tag{2.3.7}$$

Then equation (2.3.3) can be rewritten as

$$\inf_{c \in C(I)} \left\{ \frac{DV}{D\tau} - (c - a(c))P - \mathcal{L}V \right\} = 0. \tag{2.3.8}$$

Let us analyze (2.3.6-2.3.8) from a discrete point of view, that is, consider discrete grid points and discrete times. Let $\mathcal{I}\big(\tau; P_i, I_j, \tau^{n+1}, \zeta_{i,j}(\tau)\big)$ denote a path satisfying (2.3.6), which arrives at a discrete grid point (P_i, I_j) at $\tau = \tau^{n+1}$ for P_i being held constant and control following a path $\zeta_{i,j}(\tau)$ with respect to τ. Let $\mathcal{I}\big(\tau^n; P_i, I_j, \tau^{n+1}, \zeta_{i,j}(\tau^n)\big)$ be the

departure point of this path at $\tau = \tau^n$, which can be computed by solving

$$\begin{cases} \dfrac{d\mathcal{I}}{d\tau}\big(\tau; P_i, I_j, \tau^{n+1}, \zeta_{i,j}(\tau)\big) = \zeta_{i,j}(\tau) + a\big(\zeta_{i,j}(\tau)\big) & \text{for } \tau < \tau^{n+1}, \\ \\ \mathcal{I}\big(\tau; P_i, I_j, \tau^{n+1}, \zeta_{i,j}(\tau)\big) = I_j & \text{for } \tau = \tau^{n+1}, \end{cases} \tag{2.3.9}$$

from $\tau = \tau^{n+1}$ to $\tau = \tau^n$. We can write the solution of (2.3.9) in the integral form as

$$\mathcal{I}\big(\tau = \tau^n; P_i, I_j, \tau^{n+1}, \zeta_{i,j}(\tau = \tau^n)\big) = I_j - \int_{\tau^n}^{\tau^{n+1}} \Big[\zeta_{i,j}(\tau) + a\big(\zeta_{i,j}(\tau)\big)\Big] d\tau. \tag{2.3.10}$$

Note that from (2.3.10), the departure point $\mathcal{I}\big(\tau^n; P_i, I_j, \tau^{n+1}, \zeta_{i,j}(\tau^n)\big)$ will not necessarily coincide with a grid point in the I direction. To simplify the notation, in the following we will use $\mathcal{I}(\tau) = \mathcal{I}\big(\tau; P_i, I_j, \tau^{n+1}, \zeta_{i,j}(\tau)\big)$ without causing ambiguity. An example of the semi-Lagrangian trajectory $\mathcal{I}(\tau)$ is illustrated in Figure 2.1.

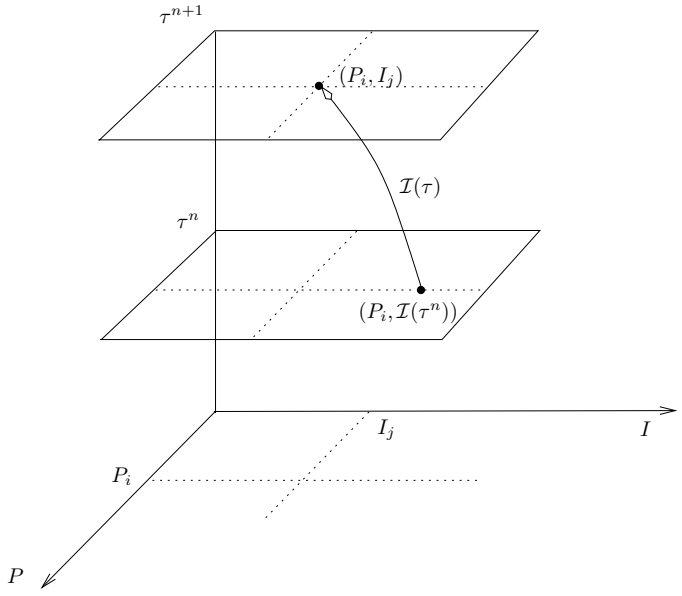


FIGURE 2.1: *Illustration of a semi-Lagrangian trajectory $\mathcal{I}(\tau)$ that arrives at a grid node $\mathcal{I}(\tau^{n+1}) = I_j$ at $\tau = \tau^{n+1}$ from the departure point $\mathcal{I}(\tau^n)$ at $\tau = \tau^n$, where the value of P remains at P_i. Note that $\mathcal{I}(\tau^n)$ normally does not correspond to a discrete grid node in the I direction.*

Integrating both sides of equation (2.3.8) along the path $\mathcal{I}(\tau)$ from $\tau = \tau^n$ to $\tau = \tau^{n+1}$,

with P fixed at P_i and control variable c following the path $\zeta_{i,j}(\tau)$, gives

$$\int_{\tau^n}^{\tau^{n+1}} \inf_{\zeta_{i,j}(\tau)\in C(\mathcal{I}(\tau))} \left\{ \frac{DV}{D\tau}\big(P_i, \mathcal{I}(\tau), \tau\big) - \big(\zeta_{i,j}(\tau) - a(\zeta_{i,j}(\tau))\big) P_i - \mathcal{L}V\big(P_i, \mathcal{I}(\tau), \tau\big) \right\} d\tau = 0. \tag{2.3.11}$$

Interchanging the integral and the inf operator in (2.3.11), assuming that they are interchangeable, and using the identity

$$\int_{\tau^n}^{\tau^{n+1}} \frac{DV}{D\tau}\big(P_i, \mathcal{I}(\tau), \tau\big) d\tau = V\big(P_i, I_j, \tau^{n+1}\big) - V\big(P_i, \mathcal{I}(\tau^n), \tau^n\big) \tag{2.3.12}$$

we obtain

$$V\big(P_i, I_j, \tau^{n+1}\big) = \sup_{\zeta_{i,j}(\tau)\in C(\mathcal{I}(\tau))} \left\{ V\big(P_i, \mathcal{I}(\tau^n), \tau^n\big) + \int_{\tau^n}^{\tau^{n+1}} \big(\zeta_{i,j}(\tau) - a\big(\zeta_{i,j}(\tau)\big)\big) P_i d\tau + \int_{\tau^n}^{\tau^{n+1}} \mathcal{L}V\big(P_i, \mathcal{I}(\tau), \tau\big) d\tau \right\}, \tag{2.3.13}$$

where $\mathcal{I}(\tau) = \mathcal{I}\big(\tau; P_i, I_j, \tau^{n+1}, \zeta_{i,j}(\tau)\big)$.

Remark 2.1 (Interchanging the Order of Operations in (2.3.11))**.** *The integral and the* inf *operator may not be interchangeable. Moreover, the derivatives in equation (2.3.11) may not exist since the value function V may not be smooth and needs to be considered in the sense of the viscosity solution. Thus, our derivation is not rigorous. However, our purpose here is to illustrate the main idea for developing the schemes. The rigorous proof of the convergence of the semi-Lagrangian fully implicit discretization to the viscosity solution of equation (2.2.12) will be given in Section 3.2.*

By evaluating the integrals in equations (2.3.10) and (2.3.13) using various numerical integration schemes, we are able to obtain semi-Lagrangian discretizations of different orders in time. In this section, we will present the fully implicit and Crank-Nicolson timestepping schemes which result from approximating the integrals using the rectangular

rule and trapezoidal rule, respectively. We will use an approach similar to that suggested in [43].

2.3.2 Fully Implicit Timestepping

In the case of fully implicit timestepping, we approximate the integral in equation (2.3.10) using the rectangular rule at $\tau = \tau^{n+1}$. In other words, we evaluate (2.3.10) by assuming that

$$\zeta_{i,j}(\tau) \simeq \zeta_{i,j}(\tau^{n+1}) \quad \text{for } \tau \in [\tau^n, \tau^{n+1}]. \tag{2.3.14}$$

It is perhaps not immediately obvious why we evaluate $\zeta_{i,j}(\tau)$ at τ^{n+1} in approximation (2.3.14). In Appendix C we show that this choice corresponds to a discretization based on assuming that the operator of the facility can adjust the controls only at finite intervals, and that no-arbitrage jump conditions are applied at the control choice times. As a consequence, the fully implicit semi-Lagrangian discretization satisfies discrete no-arbitrage jump conditions.

Let $\zeta_{i,j}^{n+1} = \zeta_{i,j}(\tau^{n+1})$ and let $I_{\hat{j}}^n$ denote an approximation to $\mathcal{I}(\tau^n) = \mathcal{I}\big(\tau^n; P_i, I_j, \tau^{n+1}, \zeta_{i,j}(\tau^n)\big)$, the departure point of the semi-Lagrangian trajectory (2.3.9). The rectangular approximation of (2.3.10), assuming (2.3.14), gives

$$I_{\hat{j}}^n = I_j - \Delta\tau\big(\zeta_{i,j}^{n+1} + a\big(\zeta_{i,j}^{n+1}\big)\big), \tag{2.3.15}$$

where $\Delta\tau = \tau^{n+1} - \tau^n$.

The control $\zeta_{i,j}^{n+1}$ must satisfy the constraint $\zeta_{i,j}^{n+1} \in C(I_j)$, where $C(I_j) = [c_{\min}(I_j), -k_5] \cup [0, c_{\max}(I_j)]$, as defined in equation (2.2.7). Moreover, to prevent the value of $I_{\hat{j}}^n$ from going outside of the domain $[0, I_{\max}]$, we need to impose further constraints on $\zeta_{i,j}^{n+1}$. Let $C_j^{n+1} \subseteq C(I_j)$ denote the set of values of $\zeta_{i,j}^{n+1}$ such that the resulting $I_{\hat{j}}^n$ calculated from equation (2.3.15) is bounded within $[0, I_{\max}]$. Note that C_j^{n+1} is independent of P_i. We regard all elements in C_j^{n+1} as admissible controls.

Equation (2.3.15) provides $I_{\hat{j}}^n$ as an approximation to $\mathcal{I}(\tau^n)$. Hence, $V\big(P_i, I_{\hat{j}}^n, \tau^n\big)$ is an approximation to $V(P_i, \mathcal{I}(\tau^n), \tau^n)$, which is the value function at τ^n when P is fixed

at P_i and I residing at the departure point of the path $\mathcal{I}(\tau)$. As mentioned above, $I^n_{\hat{j}}$ usually does not coincide with a grid point in I direction. Thus, we have to choose an interpolation scheme to approximate $V\big(P_i, I^n_{\hat{j}}, \tau^n\big)$ using discrete grid values $V^n_{i,j}$, $i = 0, \ldots, i_{\max}$, $j = 0, \ldots, j_{\max}$. Let $V^n_{i,\hat{j}}$ denote an approximation of $V\big(P_i, I^n_{\hat{j}}, \tau^n\big)$ obtained by interpolating a set of values $V^n_{i,j}$ with P fixed at P_i and I varied. Therefore, we have

$$\begin{aligned} V^n_{i,\hat{j}} &= V\big(P_i, I^n_{\hat{j}}, \tau^n\big) + \text{ Error} \\ &= V(P_i, \mathcal{I}(\tau^n), \tau^n) + \text{ Error}. \end{aligned} \tag{2.3.16}$$

Evaluating the integrals in (2.3.13) at $\tau = \tau^{n+1}$ using the rectangular rule, assuming that the control path $\zeta_{i,j}(\tau)$ follows (2.3.14) and the semi-Lagrangian trajectory $\mathcal{I}(\tau)$ satisfies (2.3.15), gives

$$V^{n+1}_{i,j} = \sup_{\zeta^{n+1}_{i,j} \in C^{n+1}_j} \Big\{ V^n_{i,\hat{j}} + \Delta\tau\big(\zeta^{n+1}_{i,j} - a\big(\zeta^{n+1}_{i,j}\big)\big) P_i \Big\} + \Delta\tau (\mathcal{L}_h V)^{n+1}_{i,j}, \tag{2.3.17}$$

where $V(P_i, \mathcal{I}(\tau^n), \tau^n)$ in (2.3.13) is approximated by $V^n_{i,\hat{j}}$. The last term in (2.3.17) follows from approximating the second integral in (2.3.13) assuming $\mathcal{L}V(P_i, \mathcal{I}(\tau), \tau) = \mathcal{L}V(P_i, \mathcal{I}(\tau^{n+1}), \tau^{n+1}) = \mathcal{L}V(P_i, I_j, \tau^{n+1})$ for $\tau \in [\tau^n, \tau^{n+1}]$ and then replacing the differential operator $\mathcal{L}V$ with its discrete form $\mathcal{L}_h V$, given in (2.3.4). Equations (2.3.15-2.3.17) specify a semi-Lagrangian fully implicit discretization. Assuming the solution value is smooth, although this is not the case in general, we show in Lemma 3.19 that linear interpolation for computing $V^n_{i,\hat{j}}$ is sufficient to achieve a first-order global discretization error. We will also demonstrate the first-order convergence of the fully implicit timestepping scheme using numerical experiments in Chapter 4.

Using an entirely different approach, in Appendix C, we derive a semi-discretization based on a discrete optimal control approximation, and no-arbitrage jump conditions. If we further discretize this method in the $(P, \mathcal{I})$ directions, we obtain a discretization which is algebraically identical to the fully implicit discretization (2.3.15-2.3.17).

2.3.3 Crank-Nicolson Timestepping

In order to obtain a higher order discretization in time, we can evaluate the integrals in (2.3.10) and (2.3.13) using a trapezoidal rule, which results in a Crank-Nicolson timestepping scheme. We assume that the control path $\zeta_{i,j}(\tau)$ is a continuous differentiable function of time. Let $\zeta_{i,j}^n = \zeta_{i,j}(\tau^n)$. Applying the trapezoidal rule to the integral in (2.3.10), assuming the control is a smooth function of time, gives the following approximation for $I_{\hat{j}}^n$

$$I_{\hat{j}}^n = I_j - \frac{\Delta\tau}{2}\big(\zeta_{i,j}^{n+1} + a\big(\zeta_{i,j}^{n+1}\big)\big) - \frac{\Delta\tau}{2}\big(\zeta_{i,j}^{n} + a\big(\zeta_{i,j}^{n}\big)\big). \tag{2.3.18}$$

Similar to the definition of admissible controls in the previous subsection, we can define $C_j^{n+1} \subseteq C\big(I_{\hat{j}}^n\big) \times C(I_j)$ be the set of all admissible controls $\zeta_{i,j}^n$ and $\zeta_{i,j}^{n+1}$ such that the value $I_{\hat{j}}^n$ calculated from (2.3.18) resides inside the domain $[0, I_{\max}]$.

Approximating the integrals in (2.3.13) using the trapezoidal rule, assuming that the control path $\zeta_{i,j}(\tau)$ is a smooth function of time, and that the semi-Lagrangian trajectory $\mathcal{I}(\tau)$ follows (2.3.18), then we obtain

$$\begin{aligned} V_{i,j}^{n+1} = \sup_{(\zeta_{i,j}^n, \zeta_{i,j}^{n+1}) \in C_j^{n+1}} \Big\{ & V_{i,\hat{j}}^n + \frac{\Delta\tau}{2}\big(\zeta_{i,j}^{n+1} - a\big(\zeta_{i,j}^{n+1}\big)\big)P_i \\ & + \frac{\Delta\tau}{2}\big(\zeta_{i,j}^{n} - a\big(\zeta_{i,j}^{n}\big)\big)P_i + \frac{\Delta\tau}{2}(\mathcal{L}_h V)_{i,\hat{j}}^n \Big\} + \frac{\Delta\tau}{2}(\mathcal{L}_h V)_{i,j}^{n+1}, \end{aligned} \tag{2.3.19}$$

where $(\mathcal{L}_h V)_{i,\hat{j}}^n$ is the evaluation of the discrete differential operator (2.3.4) at $\tau = \tau^n$ and $(P, I) = \big(P_i, I_{\hat{j}}^n\big)$ with $I_{\hat{j}}^n$ given in equation (2.3.18). Equations (2.3.18-2.3.19) result in a semi-Lagrangian Crank-Nicolson discretization.

As in the fully implicit timestepping case, we need to use interpolation to compute quantities of the form $(\cdot)_{i,\hat{j}}^n$ in (2.3.19) since $I_{\hat{j}}^n$ usually does not reside on a grid point in I direction. As suggested by [40, 14, 44] for the case when the control is a fixed constant, second-order global truncation error can be achieved if the P derivatives in $\mathcal{L}V$ are discretized using second-order accurate methods, e.g., central differencing method (see Appendix B), and quadratic interpolation is used for $(\cdot)_{i,\hat{j}}^n$. Of course, this truncation error estimate is valid only for smooth solutions. Indeed, in the numerical experiments

conducted in Chapter 4, we cannot observe second-order convergence for the Crank-Nicolson timestepping scheme with high-order interpolants.

2.3.4 Solution Algorithm

In order to solve the discrete equations (2.3.15-2.3.17) and (2.3.18-2.3.19), we formulate the equations into a linear system. We then develop an algorithm to compute the solution by iterating through the timesteps and solving the corresponding linear system at each timestep.

Before proceeding to setting up the matrix form for the discrete equations (2.3.15-2.3.17) and (2.3.18-2.3.19), let us introduce the following notation. Let V^n denote a column vector that includes all values of $V^n_{i,j}$ with the index order arranged as

$$V^n = [V^n_{0,0}, \ldots, V^n_{i_{\max},0}, \ldots, V^n_{0,j_{\max}}, \ldots, V^n_{i_{\max},j_{\max}}]'. \tag{2.3.20}$$

Here $[\ \ ]'$ denotes transpose of a vector. For future reference, assuming M is a square matrix, we denote $[MV^n]_{ij} = (MV^n)_{i,j}$, and set

$$[MV^n]_j = \left[(MV^n)_{0,j}, \ldots, (MV^n)_{i_{\max},j}\right]', \tag{2.3.21}$$

where the index of $(MV^n)_{i,j}$ in MV^n is the same as that of $V^n_{i,j}$ in V^n.

Based on the discrete differential operator $\mathcal{L}_h V$ in (2.3.4), we can define a matrix L^n such that

$$\begin{aligned}[L^n \cdot V^n]_{ij} &= (\mathcal{L}_h V)^n_{i,j} \\ &= [\gamma^n_i V^n_{i-1,j} + \beta^n_i V^n_{i+1,j} - (\gamma^n_i + \beta^n_i + r)V^n_{i,j}],\end{aligned} \tag{2.3.22}$$

where the coefficients γ^n_i and β^n_i are given in Appendix B.

Let Φ^{n+1} be a Lagrange interpolation operator such that

$$\left[\Phi^{n+1} \cdot V^n\right]_{ij} = V^n_{i,\hat{j}} + \text{interpolation error}, \tag{2.3.23}$$

where Φ^{n+1} can represent any order (linear, quadratic) of Lagrangian interpolation. Let $\left[\Phi^{n+1}V^n\right]_j$ denote a column vector with entries

$$\left[\left[\Phi^{n+1}V^n\right]_j\right]_i = \left[\Phi^{n+1}V^n\right]_{i,j} \quad . \tag{2.3.24}$$

Let P denote a column vector satisfying $[P]_i = P_i$. Let ζ_j^n and ζ_j^{n+1} be diagonal matrices with diagonal entries $\left[\zeta_j^n\right]_{ii} = \zeta_{i,j}^n$ and $\left[\zeta_j^{n+1}\right]_{ii} = \zeta_{i,j}^{n+1}$. Similarly, let $a\big(\zeta_j^n\big)$ and $a\big(\zeta_j^{n+1}\big)$ denote diagonal matrices with diagonal entries $\left[a\big(\zeta_j^n\big)\right]_{ii} = a\big(\zeta_{i,j}^n\big)$, $\left[a\big(\zeta_j^{n+1}\big)\right]_{ii} = a\big(\zeta_{i,j}^{n+1}\big)$. Let I be an identity matrix. Given the above notations, the discrete equations (2.3.15-2.3.17) and (2.3.18-2.3.19) together can be written in a θ-form as

$$\begin{aligned}
&\left[[\mathsf{I} - (1-\theta)\Delta\tau L^{n+1}]V^{n+1}\right]_j = \left[\Phi^{n+1}\left[\mathsf{I} + \theta\Delta\tau L^n\right]V^n\right]_j + \\
&\qquad\qquad (1-\theta)\Delta\tau\big(\zeta_j^{n+1} - a\big(\zeta_j^{n+1}\big)\big)P + \theta\Delta\tau\big(\zeta_j^n - a\big(\zeta_j^n\big)\big)P \\
&\text{where } \left[\zeta_j^n\right]_{ii}, \left[\zeta_j^{n+1}\right]_{ii} = \underset{([\zeta_j^n]_{ii},[\zeta_j^{n+1}]_{ii})\in C_j^{n+1}}{\arg\max} \Big\{ \Big[\left[\Phi^{n+1}[\mathsf{I} + \theta\Delta\tau L^n]V^n\right]_j + \\
&\qquad\qquad (1-\theta)\Delta\tau\big(\zeta_j^{n+1} - a\big(\zeta_j^{n+1}\big)\big)P + \theta\Delta\tau\big(\zeta_j^n - a\big(\zeta_j^n\big)\big)P\Big]_i\Big\}
\end{aligned} \tag{2.3.25}$$

for $j = 0, \ldots, j_{\max}$. Here $\theta = 0$ corresponds to fully implicit timestepping, and $\theta = 1/2$ is Crank-Nicolson timestepping. Note that we can use the operation $\arg\max$ in (2.3.25) because the supremums in (2.3.17) and (2.3.19) can be achieved by a control $\zeta_{i,j}^{n+1}$ and a control pair $(\zeta_{i,j}^n, \zeta_{i,j}^{n+1})$, respectively, according to the arguments in Section 2.4.

After setting the matrix form, the solution to HJB equation (2.2.12) and the associated boundary conditions can be computed using Algorithm 2.1.

Remark 2.2. *As described in [46], a standard implicit finite difference discretization for equation (2.2.12) requires a Policy type iteration at each timestep to solve the nonlinear discretized algebraic equations. An alternative approach uses an explicit timestepping method, but an explicit method suffers from the usual parabolic stability condition. However, Algorithm 2.1, which uses a semi-Lagrangian discretization, avoids the need for Policy iteration. Instead, Algorithm 2.1 replaces the non-linearity involving V^{n+1} with local optimization problems involving V^n, the known solution from the previous timestep.*

$V^0 = \text{Option Payoff}$
For $n = 0, \ldots,$ // Timestep loop
  For $j = 1, \ldots,$ // Loop through nodes in I direction
    For $i = 1, \ldots,$ // Loop through nodes in P direction

$$\left[\zeta_j^n\right]_{ii}, \left[\zeta_j^{n+1}\right]_{ii} = \underset{([\zeta_j^n]_{ii}, [\zeta_j^{n+1}]_{ii}) \in C_j^{n+1}}{\operatorname{argmax}} \Big\{ \Big[\left[\Phi^{n+1}[\mathsf{I} + \theta\Delta\tau L^n]V^n\right]_j + (1-\theta)\Delta\tau\big(\zeta_j^{n+1} - a\big(\zeta_j^{n+1}\big)\big)P + \theta\Delta\tau\big(\zeta_j^n - a\big(\zeta_j^n\big)\big)P\Big]_i \Big\}$$

    EndFor
    Solve

$$\left[[\mathsf{I} - (1-\theta)\Delta\tau L^{n+1}]V^{n+1}\right]_j = \left[\Phi^{n+1}\left[\mathsf{I} + \theta\Delta\tau L^n\right]V^n\right]_j + (1-\theta)\Delta\tau\big(\zeta_j^{n+1} - a\big(\zeta_j^{n+1}\big)\big)P + \theta\Delta\tau\big(\zeta_j^n - a\big(\zeta_j^n\big)\big)P$$

  EndFor
EndFor

ALGORITHM 2.1: *Semi-Lagrangian timestepping*

Remark 2.3. *At each timestep in Algorithm 2.1 all discrete equations in the I direction are decoupled and independent of each other. This property makes solution of the gas storage contract straightforward to implement.*

2.4 Solving the Local Optimization Problem

According to Algorithm 2.1, we need to solve a constrained optimization problem

$$\sup_{(\zeta_{i,j}^n, \zeta_{i,j}^{n+1}) \in C_j^{n+1}} \Big\{ V_{i,\hat{j}}^n + (1-\theta)\Delta\tau\big(\zeta_{i,j}^{n+1} - a\big(\zeta_{i,j}^{n+1}\big)\big)P_i + \theta\Delta\tau\big(\zeta_{i,j}^n - a\big(\zeta_{i,j}^n\big)\big)P_i + \theta\Delta\tau(\mathcal{L}_h V)_{i,\hat{j}}^n \Big\}, \tag{2.4.1}$$

$$\text{with } I_{\hat{j}}^n = I_j - (1-\theta)\Delta\tau\big(\zeta_{i,j}^{n+1} + a\big(\zeta_{i,j}^{n+1}\big)\big) - \theta\Delta\tau\big(\zeta_{i,j}^n + a\big(\zeta_{i,j}^n\big)\big) \tag{2.4.2}$$

at every mesh node (P_i, I_j) and at every discrete timestep τ^n.

According to [80], the exact solution of equation (2.2.12) has the property that the controls are of the *bang-bang* type, i.e. the optimal controls can take on only the values

in a finite set $\{0, c_{\max}(I), c_{\min}(I)\}$. Nevertheless, for a finite grid size, the solution of the discrete optimization problem (2.4.1) may allow controls which are not optimal controls for the exact solution of the HJB equation. Consequently, there are two possible approaches for determining the optimal controls at each grid node. We can use our knowledge of the exact controls to search only for optimal controls within the known finite set of possible values. In other words, the set of admissible controls is finite in this case, hence this approach is consistent with the control behaviour in the exact solution. We refer to this approach as the *bang-bang* method. Alternatively, we can simply solve the discrete optimization problem in (2.4.1), and allow any admissible control. We will refer to this technique as the *no bang-bang* method.

In this section, we give an overview of the bang-bang and no bang-bang methods for solving problem (2.4.1). The details are tedious and therefore are omitted in this thesis.

No bang-bang method

Problem (2.4.1) is nonlinear since the admissible set C_j^{n+1} for controls depends on the value of $I_{\hat{j}}^n$, which in turn is a function of controls. This makes it difficult to directly solve problem (2.4.1), especially in the case of Crank-Nicolson timestepping.

We simplify the problem by changing unknowns from $\zeta_{i,j}^n, \zeta_{i,j}^{n+1}$ to $I_{\hat{j}}^n$. Specifically, we can write equation (2.4.2) as

$$(1-\theta)\Delta\tau\zeta_{i,j}^{n+1} + \theta\Delta\tau\zeta_{i,j}^{n} = \big(I_j - I_{\hat{j}}^n\big) - (1-\theta)\Delta\tau a\big(\zeta_{i,j}^{n+1}\big) - \theta\Delta\tau a\big(\zeta_{i,j}^{n}\big). \tag{2.4.3}$$

Substituting equation (2.4.3) into (2.4.1) leads to

$$\begin{aligned}\sup_{(\zeta_{i,j}^n, \zeta_{i,j}^{n+1}) \in C_j^{n+1}} \Big\{ & V_{i,\hat{j}}^n - I_{\hat{j}}^n P_i + \theta\Delta\tau(\mathcal{L}_h V)_{i,\hat{j}}^n \\ & - 2(1-\theta)\Delta\tau a\big(\zeta_{i,j}^{n+1}\big)P_i - 2\theta\Delta\tau a\big(\zeta_{i,j}^{n}\big)P_i \Big\} + I_j P_i \quad .\end{aligned} \tag{2.4.4}$$

In the following we consider Crank-Nicolson timestepping ($\theta = 1/2$) in (2.4.4); the same method can be applied to fully implicit timestepping, which is a much easier problem

compared to Crank-Nicolson timestepping. Since the values of $a(\zeta_{i,j}^{n+1})$ and $a(\zeta_{i,j}^{n})$ are either 0 or k_5 according to the signs of $\zeta_{i,j}^{n+1}$ and $\zeta_{i,j}^{n}$ (see (2.2.5)), we can drop the dependence of the objective function in (2.4.4) on $\zeta_{i,j}^{n+1}$, $\zeta_{i,j}^{n}$ by separately considering four regions corresponding to the different combinations of signs of $\zeta_{i,j}^{n+1}$ and $\zeta_{i,j}^{n}$. For example, one region is $(\zeta_{i,j}^{n+1}, \zeta_{i,j}^{n}) \in [0, c_{\max}(I_j)] \times \left[0, c_{\max}\left(I_{\hat{j}}^{n}\right)\right]$; when $\zeta_{i,j}^{n+1}, \zeta_{i,j}^{n}$ lie in this region we have $a(\zeta_{i,j}^{n+1}) = a(\zeta_{i,j}^{n}) = 0$. Meanwhile, for all controls $\zeta_{i,j}^{n+1}, \zeta_{i,j}^{n}$ in any of the four regions, it can be shown that the corresponding values of $I_{\hat{j}}^{n}$ consist of a closed interval $[I_{\min}^{n}, I_{\max}^{n}] \subseteq [0, I_{\max}]$, where the bounds $I_{\min}^{n}$ and $I_{\max}^{n}$ are independent of controls $\zeta_{i,j}^{n+1}$ and $\zeta_{i,j}^{n}$. Therefore, through changing of unknowns, instead of solving one two-dimensional nonlinear optimization problem (2.4.1), identically, we can solve four one-dimensional optimization problems and pick the maximum value among the four; each of the optimization problems has the form

$$\max_{I_{\hat{j}}^{n} \in [I_{\min}^{n}, I_{\max}^{n}]} \left\{ V_{i,\hat{j}}^{n} - I_{\hat{j}}^{n} P_i + \theta \Delta\tau (\mathcal{L}_h V)_{i,\hat{j}}^{n} \right\} - 2(1-\theta)\Delta\tau D^{n+1} P_i - 2\theta\Delta\tau D^{n} P_i + I_j P_i, \tag{2.4.5}$$

where D^n, D^{n+1} are constants determined by the signs of $\zeta_{i,j}^{n}$ and $\zeta_{i,j}^{n+1}$. Note that $V_{i,\hat{j}}^{n} - I_{\hat{j}}^{n} P_i + \theta\Delta\tau(\mathcal{L}_h V)_{i,\hat{j}}^{n}$ is a function of $I_{\hat{j}}^{n}$, representing a curve constructed by linear or quadratic interpolation using discrete values $V_{i,j}^{n}$ and $(\mathcal{L}_h V)_{i,j}^{n}$, $j = 0, \ldots, j_{\max}$. The curve generated by our interpolation method (see below) is continuous and bounded. As a result, the supremum is achieved by a value of $I_{\hat{j}}^{n}$ since the interval $[I_{\min}^{n}, I_{\max}^{n}]$ is closed. Therefore, we can use the max expression to replace the sup expression.

If linear interpolation is used, then the optimal value of $I_{\hat{j}}^{n}$ for problem (2.4.5) resides either at boundaries $I_{\min}^{n}, I_{\max}^{n}$ or at discrete grid nodes lying between $I_{\min}^{n}$ and $I_{\max}^{n}$. As a result, we only need to check the boundaries and the discrete nodes and return the maximum value as the solution to problem (2.4.5).

If quadratic interpolation is used, then we divide interval $[I_{\min}^{n}, I_{\max}^{n}]$ into several sub-intervals; within each sub-interval, the same interpolation stencils are used to compute quantities $(\cdot)_{i,\hat{j}}^{n}$. For each sub-interval, we calculate the maximum value of the objective function in (2.4.5) with $I_{\hat{j}}^{n}$ residing inside that sub-interval. Finally, we compare the values

calculated for all sub-intervals and select the maximum as the solution to problem (2.4.5). The sub-intervals are chosen such that the interpolation curves are continuous within $[I^n_{\min}, I^n_{\max}]$. To reduce the effect of non-monotonicity caused by quadratic interpolation, we limit the interpolation by requiring [93, 14] (assuming $I_s \le I^n_{\hat{j}} \le I_{s+1}$)

$$V^n_{i,\hat{j}} \le \max\{V^n_{i,s}\ ,\ V^n_{i,s+1}\}\ \ ;\ \ V^n_{i,\hat{j}} \ge \min\{V^n_{i,s}\ ,\ V^n_{i,s+1}\}. \tag{2.4.6}$$

In the case of fully implicit timestepping, after obtaining the optimal $I^n_{\hat{j}}$ for problem (2.4.1), we can compute the optimal control $\zeta^{n+1}_{i,j}$ from equation (2.4.3). In this case, the variables $I^n_{\hat{j}}$ and $\zeta^{n+1}_{i,j}$ have physical meanings. Recall the discrete optimal control scenario described in Appendix C. Under this scenario, I_j and $I^n_{\hat{j}}$ represent the gas inventory at the forward times $t = t^k$ and $t = t^{k+1}$, respectively, where $t^k = T - \tau^{n+1}$; $\zeta^{n+1}_{i,j}$ represent the optimal operation that the operator imposes on the storage facility during the interval $[t^k, t^{k+1})$. In the case of Crank-Nicolson timestepping, although we can solve problem (2.4.1), given the optimal value of $I^n_{\hat{j}}$, we cannot uniquely determine the values of the control variables $\zeta^n_{i,j}$ and $\zeta^{n+1}_{i,j}$ from (2.4.3). Nor do these variables have simple physical meanings.

Bang-bang method

As shown in [80], the optimization problem

$$\sup_{c\in C(I)} \{(c - a(c))P - (c + a(c))V_I\} \tag{2.4.7}$$

in PDE (2.2.12) exhibits a bang-bang control feature. Specifically, the optimal value for c in (2.4.7) is either $c_{\max}(I)$, $c_{\min}(I)$, or 0. Therefore, we can reformulate (2.4.7) into an equivalent equation given as follows:

$$\sup_{c\in\{c_{\max}(I), c_{\min}(I), 0\}} \{(c - a(c))P - (c + a(c))V_I\}. \tag{2.4.8}$$

The bang-bang method will solve the discrete optimization problem resulting from discretizing the following gas pricing PDE equivalent to (2.2.12):

$$V_\tau = \frac{1}{2}(\hat{\sigma}(P)P)^2 V_{PP} + \alpha(K(t)-P)V_P + \sup_{c\in\{c_{\max}(I),c_{\min}(I),0\}} \{(c-a(c))P - (c+a(c))V_I\} - rV. \tag{2.4.9}$$

Since (2.2.12) and (2.4.9) have the same viscosity solution, the maximums determined with the bang-bang and no bang-bang methods will coincide as $h \to 0$.

To obtain the optimization problem corresponding to PDE (2.4.9), we can force the controls in (2.4.1) to be of bang-bang type by only examining controls $\zeta_{i,j}^{n+1}, \zeta_{i,j}^{n}$ that satisfy

$$\zeta_{i,j}^{n+1} \in \left\{c_{\max}(I_j), c_{\min}(I_j), 0\right\} \;\; ; \;\; \zeta_{i,j}^{n} \in \left\{c_{\max}(I_{\hat{j}}^n), c_{\min}(I_{\hat{j}}^n), 0\right\} . \tag{2.4.10}$$

Note that the controls from (2.4.10) may not be admissible[1], that is, the resulting $I_{\hat{j}}^n$ calculated from (2.4.2) either lies outside of domain $[0, I_{\max}]$ or does not exist.

Taking the admissible control requirement into consideration, our bang-bang control method is summarized as follows: for each pair $(\zeta_{i,j}^{n+1}, \zeta_{i,j}^{n})$ from (2.4.10), if it is admissible, then we evaluate the objective function in (2.4.1) using the pair and save the evaluated value as a candidate solution for problem (2.4.1).

Otherwise, assume that a pair $(\zeta_{i,j}^{n+1}, \zeta_{i,j}^{n})$ satisfying (2.4.10) is not admissible, we evaluate the objective function in (2.4.1) using admissible controls that reside in the same region as the pair $(\zeta_{i,j}^{n+1}, \zeta_{i,j}^{n})$ and result in the maximum or minimum bounds of $I_{\hat{j}}^n$. To illustrate the idea, let us consider a specific case when $(\zeta_{i,j}^{n+1}, \zeta_{i,j}^{n}) = \left(c_{\max}(I_j), c_{\max}\left(I_{\hat{j}}^n\right)\right)$ and $(\zeta_{i,j}^{n+1}, \zeta_{i,j}^{n})$ is not admissible. In this case, the pair $(\zeta_{i,j}^{n+1}, \zeta_{i,j}^{n})$ resides in region $[0, c_{\max}(I_j)] \times [0, c_{\max}\left(I_{\hat{j}}^n\right)]$. As explained above in the no bang-bang method, for all controls in this region, the corresponding values of $I_{\hat{j}}^n$ consist of an interval $[I_{\min}^n, I_{\max}^n]$. Then we compute the value of the objective function in (2.4.1) by using two pairs of admissible controls that result in $I_{\hat{j}}^n = I_{\min}^n$ and $I_{\hat{j}}^n = I_{\max}^n$, respectively. A similar strategy is applied if other pairs of controls satisfying (2.4.10) are not admissible.

After evaluating the objective function in (2.4.1) using different pairs of admissible

[1] However, according to Lemma D.1, $(\zeta_{i,j}^{n+1}, \zeta_{i,j}^{n})$ are always admissible if $\Delta\tau$ is sufficiently small.

controls, as described above, we return the maximum value as the solution to problem (2.4.1).

Remark 2.4. *The maximum determined with the no bang-bang approach is consistent with the exact maximum of (2.4.1) for any smooth test function, with numerical error bounded by $O(h^2)$, where h is the mesh size/timestep parameter given in (2.3.1).*

On the other hand, the maximum determined with the bang-bang approach is consistent with the exact maximum of the local optimization problem corresponding to the HJB equation (2.4.9).

Remark 2.5. *When incorporating nonlinear revenue structures, the control for the resulting equation is not of bang-bang type. The no bang-bang method (but not the bang-bang method) can still be used to solve the optimization problem in the case of fully implicit timestepping ($\theta = 0$). Thus, the fully implicit semi-Lagrangian discretization is applicable to a wide range of HJB equations including those that inherit the no bang-bang control feature.*

Remark 2.6 (Computational Complexity)**.** *Since the controls are bounded, it is easy to see, from equations (2.3.6), (2.3.18) and (2.3.1), that the number of nodes that must be examined to solve the local optimization problem at each node, using both bang-bang and no bang-bang methods, is a constant independent of h. Since equation (2.3.6) is independent of P, we can precompute and store interpolation indices and weights. In addition, the linear system solve has been reduced to a set of decoupled one-dimensional problems. Hence the complexity per timestep of the implicit semi-Lagrangian scheme is linear in the total number of nodes. Thus the complexity per timestep of the fully implicit semi-Lagrangian scheme is the same as an explicit method, but it has the obvious advantage of being unconditionally stable.*

2.5 Summary

Our contributions in this chapter are summarized as follows:

- We formulate the valuation of gas storage facilities as a bounded stochastic control problem and heuristically derive an HJB PDE corresponding to the problem.

- We develop a fully implicit and a Crank-Nicolson finite difference schemes based on a semi-Lagrangian method for solving the HJB equation and obtain the optimal control strategies.

- We show (in Appendix C) that the fully implicit, semi-Lagrangian discretization is algebraically identical to a discretization based on assuming that the operations on the storage are performed only at discrete times.

- We show that, compared to the standard implicit methods, semi-Lagrangian timestepping methods avoid the need for Policy type iterations at each node at each timestep. Instead, the methods require solution of a discrete local optimization problem at each mesh point in order to determine the optimal control value. The optimization problem can be solved efficiently so that the complexity per timestep of the fully implicit semi-Lagrangian scheme is the same as that of an explicit method or the complexity per iteration per timestep of a standard implicit method.

Chapter 3

Convergence Analysis

In the previous chapter we have developed numerical schemes for pricing the natural gas storage PDE, corresponding to a bounded stochastic control problem. Since there is no guarantee that the classical solutions exist, the convergence analysis for the numerical schemes will be conducted within the framework of viscosity solutions. Consequently, in this chapter we give an introduction to the notation of viscosity solutions and prove the convergence of the fully implicit, semi-Lagrangian scheme to the viscosity solution. We will present some numerical results in the next chapter.

3.1 Viscosity Solution

The pricing PDE (2.2.12) is nonlinear. This implies that the PDE has in general no classical solutions; in other words, the first and second order derivatives of solution V may not exist. Consequently, it is impossible to analyze the solution using the classical approach. To solve this type of problems, the notion of 'weak' solutions—viscosity solutions—was introduced in [30] for first-order Hamilton-Jacobi equations and in [65] for the second-order HJB equations. The theory of viscosity solutions is powerful because it does not require the existence of derivatives of the solution and even allows the solution to be discontinuous (see Section 3.1.3). See [29] for a complete presentation of the viscosity solutions. Furthermore, we prefer the viscosity solution because in general it is identical

to the value function of the corresponding stochastic control problem, as discussed in Section 2.2.4.

This section defines viscosity solutions for the gas storage pricing PDE (2.2.12) and the associated boundary conditions. The definition can be easily generalized to other HJB equations or HJB variational inequalities presented in the rest of the thesis.

3.1.1 Intuition

Similar to the explanation in [6, 89], we give an intuitive introduction to viscosity solutions by studying its connection with the smooth classical solutions.

Let us define a vector $\mathbf{x} = (P, I, \tau)$, and let $DV(\mathbf{x}) = (V_P, V_I, V_\tau)$ and $D^2V(\mathbf{x}) = V_{PP}$. We can write equation (2.2.12) as

$$\begin{aligned} F\big(D^2V(\mathbf{x}), DV(\mathbf{x}), V(\mathbf{x}), \mathbf{x}\big) &\equiv V_\tau - \mathcal{L}V - \sup_{c \in C(I)} \big[(c - a(c))P - (c + a(c))V_I\big] \\ &= 0, \end{aligned} \tag{3.1.1}$$

where the operator $\mathcal{L}V$ is defined in (2.3.2)

$$\mathcal{L}V = \frac{1}{2}\hat{\sigma}(P)^2 P^2 V_{PP} + \alpha(K(t) - P)V_P - rV. \tag{3.1.2}$$

Let $\bar{\Omega} = [0, P_{\max}] \times [0, I_{\max}] \times [0, T]$ denote the closed domain where our problem is defined. Let $\Omega = [0, P_{\max}] \times [0, I_{\max}] \times (0, T]$ be the region obtained from excluding the $\tau = 0$ plane from $\bar{\Omega}$. Let $\Omega_0 = \bar{\Omega} \backslash \Omega = [0, P_{\max}] \times [0, I_{\max}] \times \{0\}$. It can be verified that the function $F(M, p, g, y)(M = D^2V, p = DV, g = V, y = \mathbf{x})$ satisfies the ellipticity condition

$$F(M, p, g, y) \leq F(N, p, g, y) \quad \text{if } M \geq N, \tag{3.1.3}$$

since in our case $\hat{\sigma}^2(P)P^2 \geq 0$. It can also be verified that $F(M, p, g, y)$ is continuous on Ω. Suppose for the moment that a smooth solution to equation (3.1.1) exists so that DV and D^2V are well defined.

Let us consider a set of smooth test functions $\phi(\mathbf{x})$. Assume there exists a single point $\mathbf{x}_0 \in \Omega$ such that $\mathbf{x}_0$ is a global maximum of $(V - \phi)$ in Ω and satisfies $V(\mathbf{x}_0) = \phi(\mathbf{x}_0)$. In other words, we have

$$\begin{aligned} V - \phi \leq 0, \quad \text{for any } \mathbf{x} \in \Omega, \\ \max(V - \phi) = V(\mathbf{x}_0) - \phi(\mathbf{x}_0) = 0, \quad \mathbf{x}_0 \in \Omega. \end{aligned} \tag{3.1.4}$$

Consequently, at $\mathbf{x}_0$, we have

$$D\phi(\mathbf{x}_0) = DV(\mathbf{x}_0) \ \text{ and } \ D^2\phi(\mathbf{x}_0) \geq D^2V(\mathbf{x}_0). \tag{3.1.5}$$

Hence, from equations (3.1.3) and (3.1.5) we have

$$\begin{aligned} F\big(D^2\phi(\mathbf{x}_0), D\phi(\mathbf{x}_0), V(\mathbf{x}_0), \mathbf{x}_0\big) &= F\big(D^2\phi(\mathbf{x}_0), DV(\mathbf{x}_0), V(\mathbf{x}_0), \mathbf{x}_0\big) \\ &\leq F\big(D^2V(\mathbf{x}_0), DV(\mathbf{x}_0), V(\mathbf{x}_0), \mathbf{x}_0\big) \\ &= 0. \end{aligned} \tag{3.1.6}$$

From the analysis above, we can observe that the equation

$$F\big(D^2V(\mathbf{x}), DV(\mathbf{x}), V(\mathbf{x}), \mathbf{x}\big) = 0 \tag{3.1.7}$$

implies the inequality

$$F\big(D^2\phi(\mathbf{x}_0), D\phi(\mathbf{x}_0), V(\mathbf{x}_0), \mathbf{x}_0\big) \leq 0 \tag{3.1.8}$$

for any smooth test function ϕ satisfying

$$\begin{aligned} V - \phi \leq 0, \quad \text{for any } \mathbf{x} \in \Omega, \\ \max(V - \phi) = V(\mathbf{x}_0) - \phi(\mathbf{x}_0) = 0, \quad \mathbf{x}_0 \in \Omega. \end{aligned} \tag{3.1.9}$$

Let us now consider another set of smooth test functions $\chi(\mathbf{x})$ such that

$$\begin{aligned} V - \chi \geq 0, \quad \text{for any } \mathbf{x} \in \Omega, \\ \min(V - \chi) = V(\mathbf{x}_0) - \chi(\mathbf{x}_0) = 0, \quad \mathbf{x}_0 \in \Omega. \end{aligned} \tag{3.1.10}$$

That is, there exists a single point $\mathbf{x}_0 \in \Omega$ such that $\mathbf{x}_0$ is a global minimum of $(V - \chi)$ in Ω and satisfies $V(\mathbf{x}_0) = \chi(\mathbf{x}_0)$. Following a similar analysis, for such test functions χ, we can obtain the inequality

$$F\big(D^2\chi(\mathbf{x}_0), D\chi(\mathbf{x}_0), V(\mathbf{x}_0), \mathbf{x}_0\big) \geq 0. \tag{3.1.11}$$

Now that we have shown that equation (3.1.1) implies inequalities (3.1.8) and (3.1.11) defined using smooth test functions, we can prove that the reverse is also true, i.e., we can derive equation (3.1.1) from inequalities (3.1.8) and (3.1.11). To show this, we can take $\phi = V$ and $\chi = V$ as the special test functions and using the arguments similar to the above to show that $F\big(D^2V(\mathbf{x}), DV(\mathbf{x}), V(\mathbf{x}), \mathbf{x}\big)$ is both non-positive and non-negative at any point $\mathbf{x}_0 \in \Omega$ since any $\mathbf{x}_0$ is both a global maximum and minimum point of $V - V$.

Therefore, assuming the existence of the classical solution for equation (3.1.1), we can obtain an equivalent specification through the inequalities (3.1.8) and (3.1.11) based on smooth test functions. However, in case the smooth solution does not exist for (3.1.1), we can still use conditions (3.1.8-3.1.9) and (3.1.10-3.1.11) to define a solution to equation (3.1.1) since all derivatives still apply to smooth test functions. This is the intuition of introducing viscosity solutions. Informally, a viscosity solution V (assuming it is continuous for the moment) to equation (2.2.12) is defined such that

- For any smooth test function ϕ with $D\phi(\mathbf{x})$ and $D^2\phi(\mathbf{x})$ well defined in Ω and

$$V - \phi \leq 0 \text{ for any } \mathbf{x} \in \Omega, \qquad V(\mathbf{x}_0) = \phi(\mathbf{x}_0), \quad \mathbf{x}_0 \in \Omega \tag{3.1.12}$$

  (ϕ touches V from above at the single point $\mathbf{x}_0$), then

$$F\big(D^2\phi(\mathbf{x}_0), D\phi(\mathbf{x}_0), V(\mathbf{x}_0), \mathbf{x}_0\big) \leq 0. \tag{3.1.13}$$

- For any smooth test function χ with $D\chi(\mathbf{x})$ and $D^2\chi(\mathbf{x})$ well defined in Ω and

$$V - \chi \geq 0 \text{ for any } \mathbf{x} \in \Omega, \qquad V(\mathbf{x}_0) = \chi(\mathbf{x}_0), \quad \mathbf{x}_0 \in \Omega \tag{3.1.14}$$

(χ touches V from below at the single point $\mathbf{x}_0$), then

$$F\big(D^2\chi(\mathbf{x}_0), D\chi(\mathbf{x}_0), V(\mathbf{x}_0), \mathbf{x}_0\big) \geq 0. \tag{3.1.15}$$

In Figure 3.1, we illustrate a synthetic non-smooth viscosity solution as well as an example of test functions.

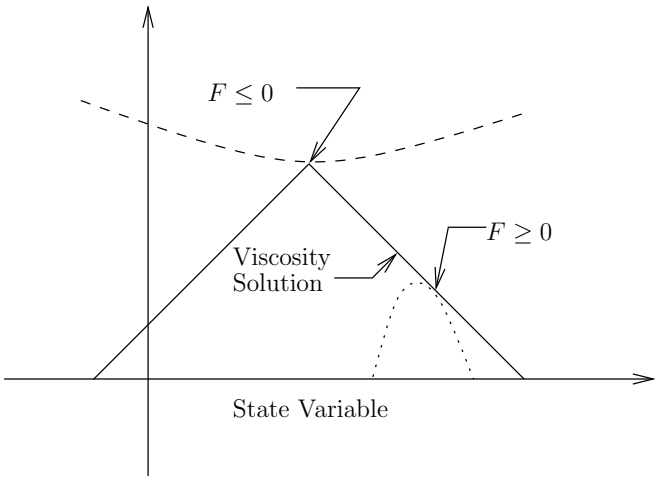


FIGURE 3.1: *Illustration of continuous viscosity solution definition. The upper dashed curve represents a smooth test function ϕ that touches the viscosity solution from above at a point $\mathbf{x}_0$, and $F\big(D^2\phi(\mathbf{x}_0), D\phi(\mathbf{x}_0), V(\mathbf{x}_0), \mathbf{x}_0\big) \leq 0$. The lower dashed curve represents a smooth test function χ that touches the viscosity solution from below at a point $\mathbf{x}_0$, and $F\big(D^2\chi(\mathbf{x}_0), D\chi(\mathbf{x}_0), V(\mathbf{x}_0), \mathbf{x}_0\big) \geq 0$. Note that there may exist some points where a smooth test function can touch the viscosity solution only from above or below, or neither. The non-smooth kink point of the viscosity solution in the figure is an example of such a point where the test function can touch only from above.*

3.1.2 Incorporating Boundary Conditions

The pricing PDE (2.2.12) is coupled with boundary conditions (2.2.13-2.2.17). Consequently, the notation of viscosity solutions must take into account the boundary behaviour. However, as pointed out in [6, 72], if a boundary equation is degenerate in the sense that the diffusion term is missing from the equation, then the solution may be discontinuous on the boundary since there is no diffusion effect to smooth out the solution across the boundary. In other words, the solution in the interior domain can be different from the solution on the boundary. Therefore, the definition of viscosity solutions needs to handle

the discontinuity of the solution on the boundary. In this section, we introduce continuous viscosity solutions that incorporate the boundary conditions. In the next section we will relax the continuity assumption and define the discontinuous viscosity solutions to handle the most general case.

Using the notation in (3.1.1), we can rewrite the pricing PDE (2.2.12) and the associated boundary conditions (2.2.13-2.2.17) in the closed domain $\bar{\Omega}$ as

$$F\big(D^2V(\mathbf{x}), DV(\mathbf{x}), V(\mathbf{x}), \mathbf{x}\big) = 0 \quad \text{if } \mathbf{x} \in \Omega = [0, P_{\max}] \times [0, I_{\max}] \times (0, T], \tag{3.1.16}$$

$$V(\mathbf{x}) - B(\mathbf{x}) = 0 \quad \text{if } \mathbf{x} \in \Omega_0 = [0, P_{\max}] \times [0, I_{\max}] \times \{0\}, \tag{3.1.17}$$

where $B(\mathbf{x})$ is a function of (P, I) representing the payoff at $\tau = 0$ and function F is defined in (3.1.1). Note that as shown in Section 2.2.5, the equations on the boundaries $I = \{0, I_{\max}\}$ and $P = \{0, P_{\max}\}$ are obtained from taking the limit of equation $F\big(D^2V(\mathbf{x}), DV(\mathbf{x}), V(\mathbf{x}), \mathbf{x}\big) = 0$ from the interior domain towards the boundaries. As a result, we use a single equation in (3.1.16) to incorporate all these equations since they are essentially the same.

Since the solution of problem (3.1.16-3.1.17) may be discontinuous on the boundary $\tau = 0$, we need to relax the condition at $\tau = 0$ in the following manner:

$$F\big(D^2V(\mathbf{x}), DV(\mathbf{x}), V(\mathbf{x}), \mathbf{x}\big) = 0 \text{ or } V(\mathbf{x}) - B(\mathbf{x}) = 0 \quad \text{if } \mathbf{x} \in \Omega_0. \tag{3.1.18}$$

This implies that the solution at $\tau = 0$ will either satisfy the equation for the interior domain or the payoff $V - B = 0$.

Using the above formulation can make it easy to present the problem. For example, at the strike price, the value of a digital option is discontinuous when $t \to T$, where T is the maturity time; while away from the strike price, the values are continuous when $t \to T$. As another example, a controlled PDE may have inward or outward characteristics for different regions on the boundary depending on the values of optimal control at each region. This will result in the solution being continuous at some regions on the boundary and being discontinuous at the other regions on the boundary. Using the formulation

(3.1.18), we do not need to enumerate all possible continuous/discontinuous regions and hence have a simple statement of the problem.

Note that condition (3.1.18) needs to be satisfied in the *viscosity sense*, i.e., the viscosity solution V (assuming it is continuous for the moment) at the boundary $\tau = 0$ is defined such that

- For any smooth test function ϕ with $D\phi(\mathbf{x})$ and $D^2\phi(\mathbf{x})$ well defined in Ω_0 and

$$V - \phi \leq 0 \text{ for any } \mathbf{x} \in \Omega_0, \qquad V(\mathbf{x}_0) = \phi(\mathbf{x}_0), \quad \mathbf{x}_0 \in \Omega_0 \tag{3.1.19}$$

(ϕ touches V from above at the single point $\mathbf{x}_0$), then

$$\begin{aligned} &F\big(D^2\phi(\mathbf{x}_0), D\phi(\mathbf{x}_0), V(\mathbf{x}_0), \mathbf{x}_0\big) \leq 0 \;\text{ or }\; V(\mathbf{x}_0) - B(\mathbf{x}_0) \leq 0 \\ \Longleftrightarrow\; &\min\big[F\big(D^2\phi(\mathbf{x}_0), D\phi(\mathbf{x}_0), V(\mathbf{x}_0), \mathbf{x}_0\big),\; V(\mathbf{x}_0) - B(\mathbf{x}_0)\big] \leq 0. \end{aligned} \tag{3.1.20}$$

- For any smooth test function χ with $D\chi(\mathbf{x})$ and $D^2\chi(\mathbf{x})$ well defined in Ω_0 and

$$V - \chi \geq 0 \text{ for any } \mathbf{x} \in \Omega_0, \qquad V(\mathbf{x}_0) = \chi(\mathbf{x}_0), \quad \mathbf{x}_0 \in \Omega_0 \tag{3.1.21}$$

(χ touches V from below at the single point $\mathbf{x}_0$), then

$$\begin{aligned} &F\big(D^2\chi(\mathbf{x}_0), D\chi(\mathbf{x}_0), V(\mathbf{x}_0), \mathbf{x}_0\big) \geq 0 \;\text{ or }\; V(\mathbf{x}_0) - B(\mathbf{x}_0) \geq 0 \\ \Longleftrightarrow\; &\max\big[F\big(D^2\chi(\mathbf{x}_0), D\chi(\mathbf{x}_0), V(\mathbf{x}_0), \mathbf{x}_0\big),\; V(\mathbf{x}_0) - B(\mathbf{x}_0)\big] \geq 0. \end{aligned} \tag{3.1.22}$$

Let us now introduce definition of the continuous viscosity solution in whole domain $\bar{\Omega} = \Omega \cup \Omega_0$. First we define functions F_- and F_+ in $\bar{\Omega}$ satisfying

$$\begin{aligned} &F_-\big(M, p, g, y\big)(M = D^2V, p = DV, g = V, y = \mathbf{x}) \\ &= \begin{cases} F\big(M, p, g, y\big) & \text{if } y \in \Omega, \\ \min\big[F\big(M, p, g, y\big), g - B(y)\big] & \text{if } y \in \Omega_0. \end{cases} \end{aligned} \tag{3.1.23}$$

and

$$F_+\big(M,p,g,y\big)(M=D^2V,p=DV,g=V,y=\mathbf{x})$$
$$=\begin{cases} F\big(M,p,g,y\big) & \text{if } y\in\Omega, \\ \max\big[F\big(M,p,g,y\big),g-B(y)\big] & \text{if } y\in\Omega_0. \end{cases} \tag{3.1.24}$$

Using the notation of (3.1.23-3.1.24) and from the discussion above, we can define a continuous viscosity solution V in $\bar{\Omega}$ as follows:

Definition 3.1 (Continuous Viscosity Solutions)**.** A continuous function $V(\mathbf{x})$ is a viscosity solution of (3.1.16-3.1.17) in the closed domain $\bar{\Omega}=[0,P_{\max}]\times[0,I_{\max}]\times[0,T]$ if

- For any smooth test function $\phi(\mathbf{x})$, $\mathbf{x}\in\bar{\Omega}$, with $D\phi(\mathbf{x})$ and $D^2\phi(\mathbf{x})$ well defined in $\bar{\Omega}$ and

$$V-\phi\le 0 \text{ for any } \mathbf{x}\in\bar{\Omega}, \qquad V(\mathbf{x}_0)=\phi(\mathbf{x}_0), \quad \mathbf{x}_0\in\bar{\Omega} \tag{3.1.25}$$

  (ϕ touches V from above at the single point $\mathbf{x}_0$),

$$F_-\big(D^2\phi(\mathbf{x}_0),D\phi(\mathbf{x}_0),V(\mathbf{x}_0),\mathbf{x}_0\big)\le 0. \tag{3.1.26}$$

- For any smooth test function $\chi(\mathbf{x})$, $\mathbf{x}\in\bar{\Omega}$, with $D\chi(\mathbf{x})$ and $D^2\chi(\mathbf{x})$ well defined in $\bar{\Omega}$ and

$$V-\chi\ge 0 \text{ for any } \mathbf{x}\in\bar{\Omega}, \qquad V(\mathbf{x}_0)=\chi(\mathbf{x}_0), \quad \mathbf{x}_0\in\bar{\Omega} \tag{3.1.27}$$

  (χ touches V from below at the single point $\mathbf{x}_0$),

$$F_+\big(D^2\chi(\mathbf{x}_0),D\chi(\mathbf{x}_0),V(\mathbf{x}_0),\mathbf{x}_0\big)\ge 0. \tag{3.1.28}$$

Remark 3.2. *In Definition 3.1, there may exist points $\mathbf{x}_0$ where none of the test functions $\phi(\mathbf{x})$, $\chi(\mathbf{x})$ exist.*

3.1.3 Discontinuous Viscosity Solutions

In Definition 3.1, the viscosity solution to the system is assumed to be continuous. In order to define discontinuous viscosity solutions, we will need to use semi-continuous functions, as introduced in Definition 3.3.

Definition 3.3 (Semi-Continuous Functions)**.** Assume X is a subset of $\mathbb{R}^N$ and $f(x) : X \to \mathbb{R}$ is a function of x defined in X. Then f is upper semi-continuous (usc) at $x_0 \in X$ if

$$\limsup_{\substack{x \to x_0 \\ x \in X}} f(x) \leq f(x_0). \tag{3.1.29}$$

A function $g(x) : X \to \mathbb{R}$ is lower semi-continuous (lsc) at $x_0 \in X$ if

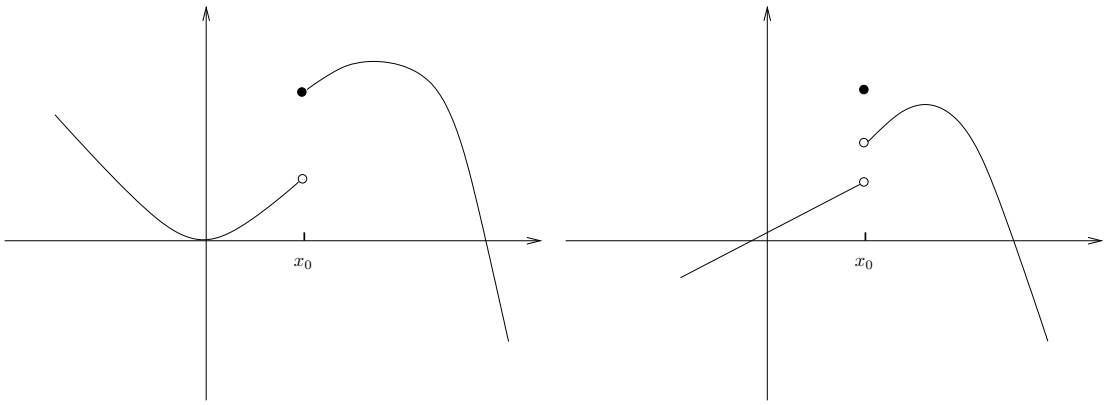
$$\liminf_{\substack{x \to x_0 \\ x \in X}} g(x) \geq g(x_0). \tag{3.1.30}$$

Remark 3.4. *If (3.1.29) holds with equality, then the function $f(x)$ will either be continuous or be left or right continuous at x_0. Figure 3.2a gives such an example. If, on the other hand, (3.1.29) holds with strict inequality, then $f(x)$ is neither left nor right continuous at x_0, and $f(x_0)$ will be strictly greater than the values of the neighbour points of x_0. An example is shown in Figure 3.2b.*

Similarly, if (3.1.30) holds with equality, then the function $g(x)$ will either be continuous or be left or right continuous at x_0. Figure 3.3a gives such an example. If (3.1.30) holds with strict inequality, then $g(x)$ is neither left nor right continuous at x_0, and $g(x_0)$ will be strictly smaller than the values of the neighbour points of x_0. An example is shown in Figure 3.3b.

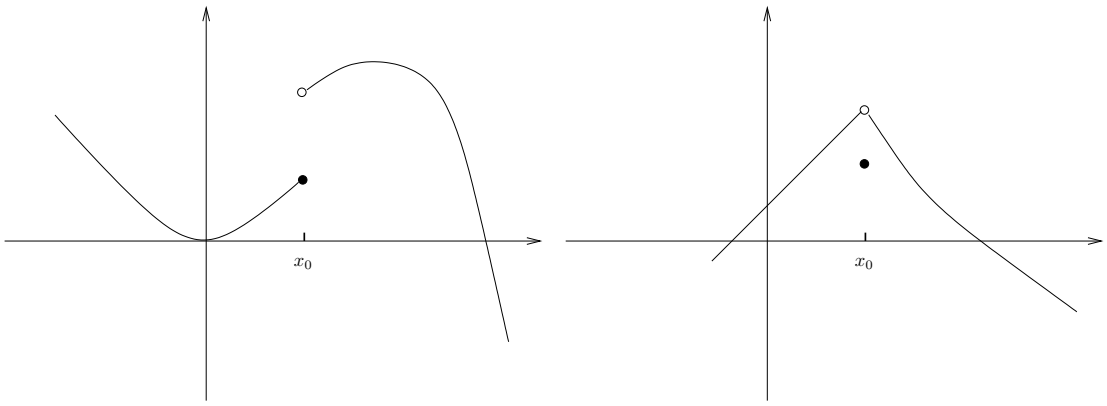
Definition 3.5 (Viscosity Subsolutions)**.** Let $f(\mathbf{x})$ be a locally bounded function defined in $\bar{\Omega}$. f is a viscosity subsolution of (3.1.16-3.1.17) if it is a usc function and if for any smooth test function $\phi(\mathbf{x})$, $\mathbf{x} \in \bar{\Omega}$, with $D\phi(\mathbf{x})$ and $D^2\phi(\mathbf{x})$ well defined in $\bar{\Omega}$ and

$$f - \phi \leq 0 \text{ for any } \mathbf{x} \in \bar{\Omega}, \qquad f(\mathbf{x}_0) = \phi(\mathbf{x}_0), \quad \mathbf{x}_0 \in \bar{\Omega} \tag{3.1.31}$$



(A) Continuous at one side (B) Discontinuous at both sides

FIGURE 3.2: *Examples of usc functions. The value of the filled point indicates* $f(x_0)$.



(A) Continuous at one side (B) Discontinuous at both sides

FIGURE 3.3: *Examples of lsc functions. The value of the filled point indicates* $g(x_0)$.

(ϕ touches f from above at the single point $\mathbf{x}_0$), we have

$$F_-\big(D^2\phi(\mathbf{x}_0), D\phi(\mathbf{x}_0), f(\mathbf{x}_0), \mathbf{x}_0\big) \leq 0. \tag{3.1.32}$$

Definition 3.6 (Viscosity Supersolutions)**.** Let $g(\mathbf{x})$ be a locally bounded function defined in $\bar{\Omega}$. g is a viscosity supersolution of (3.1.16-3.1.17) if it is a lsc function and if for any smooth test function $\chi(\mathbf{x})$, $\mathbf{x} \in \bar{\Omega}$, with $D\chi(\mathbf{x})$ and $D^2\chi(\mathbf{x})$ well defined in $\bar{\Omega}$ and

$$g - \chi \geq 0 \text{ for any } \mathbf{x} \in \bar{\Omega}, \qquad g(\mathbf{x}_0) = \chi(\mathbf{x}_0), \quad \mathbf{x}_0 \in \bar{\Omega} \tag{3.1.33}$$

(χ touches g from below at the single point $\mathbf{x}_0$), we have

$$F_+\big(D^2\chi(\mathbf{x}_0), D\chi(\mathbf{x}_0), g(\mathbf{x}_0), \mathbf{x}_0\big) \geq 0. \tag{3.1.34}$$

We can now use viscosity subsolutions and supersolutions to define discontinuous viscosity solutions. Prior to that, we need to further introduce the usc and lsc envelopes.

Definition 3.7. If C is a *closed* subset of $\mathbb{R}^N$ and $f(x) : C \to \mathbb{R}$ is a function of x defined in C, then the upper semi-continuous (usc) envelope $f^*(x) : C \to \mathbb{R}$ and the lower semi-continuous (lsc) envelope $f_*(x) : C \to \mathbb{R}$ are defined by

$$f^*(x) = \limsup_{\substack{y \to x \\ y \in C}} f(y) \quad \text{and} \quad f_*(x) = \liminf_{\substack{y \to x \\ y \in C}} f(y), \tag{3.1.35}$$

respectively. Note that in contrast to the definition of limit where only neighbour points of x excluding x itself are considered, in this definition, y in (3.1.35) includes both the point x and its neighbour points. f^* is usc in C and f_* is lsc in C.

Remark 3.8. *According to (3.1.35), we have*

$$f_*(x_0) \leq f^*(x_0), \qquad \forall x_0 \in C \tag{3.1.36}$$

and f is continuous at x_0 if (3.1.36) holds with equality.

Remark 3.9. *If f is usc in some region $C_0 \subseteq C$, then*

$$f^*(x) = f(x), \qquad \forall x \in C_0. \tag{3.1.37}$$

Similarly, if f is lsc in some region $C_1 \subseteq C$, then

$$f_*(x) = f(x), \qquad \forall x \in C_1. \tag{3.1.38}$$

Remark 3.10. *Using the usc and lsc envelopes, we can unify F_- and F^+ in (3.1.23-3.1.24) through a single equation as follows. Let $\bar{F}$ denote a function defined in $\bar{\Omega}$ given by*

$$\begin{aligned} &\bar{F}\big(M,p,g,y\big)(M = D^2V, p = DV, g = V, y = \mathbf{x}) \\ &= \begin{cases} F\big(M,p,g,y\big) & \text{if } y \in \Omega = [0, P_{\max}] \times [0, I_{\max}] \times (0,T], \\ g - B(y) & \text{if } y \in \Omega_0 = [0, P_{\max}] \times [0, I_{\max}] \times \{0\}, \end{cases} \end{aligned} \tag{3.1.39}$$

where function F is defined in (3.1.1) and $g - B(y)$ is the boundary equation at the payoff time $\tau = 0$. Then it can be verified that

$$\begin{aligned} F_- &= \bar{F}_*, \\ F_+ &= \bar{F}^* \end{aligned} \tag{3.1.40}$$

since functions $F(M,p,g,y)$ and $g - B(y)$ are continuous. Therefore, we can replace (3.1.32) and (3.1.34) with

$$\bar{F}_*\big(D^2\phi(\mathbf{x}_0), D\phi(\mathbf{x}_0), f(\mathbf{x}_0), \mathbf{x}_0\big) \leq 0. \tag{3.1.41}$$

and

$$\bar{F}^*\big(D^2\chi(\mathbf{x}_0), D\chi(\mathbf{x}_0), g(\mathbf{x}_0), \mathbf{x}_0\big) \geq 0. \tag{3.1.42}$$

We will use conditions (3.1.41) and (3.1.42) in the rest of this thesis.

We can now introduce the discontinuous viscosity solutions.

Definition 3.11 (Discontinuous Viscosity Solutions)**.** A locally bounded (possible discontinuous) function $V(\mathbf{x})$, $\mathbf{x} \in \bar{\Omega}$ is a viscosity solution of system (3.1.16-3.1.17) if its usc envelope V^* and its lsc envelope V_* are the viscosity subsolution and supersolution of (3.1.16-3.1.17), respectively.

Remark 3.12 (Non-uniqueness of the Solution)**.** *Figure 3.4 illustrates a synthetic discontinuous viscosity solution as well as the corresponding subsolution and supersolution. From the figure, we can observe an important issue: the viscosity solutions are not unique at discontinuous points. For example, in Figure 3.4a, we can reset the solution value V at x_0 as any value between $\lim_{x\to[x_0]^-} V(x)$ and $\lim_{x\to[x_0]^+} V(x)$ (the values with respect to the two unfilled balls). The resulting solution is still a valid viscosity solution since the corresponding usc and lsc envelopes remain the same.*

3.1.4 Strong Comparison Result

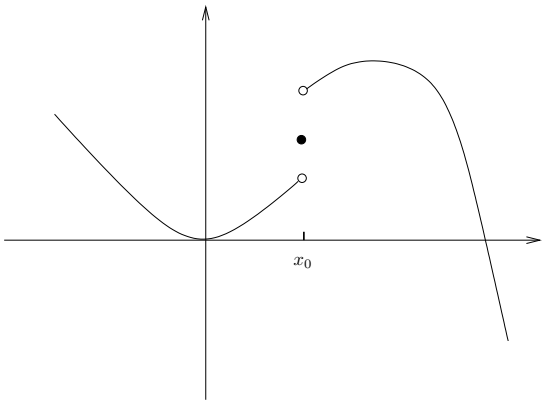
In Section 3.1.3, we give the notation for (possible) discontinuous viscosity solutions. We also show in Remark 3.12 that the viscosity solution is not unique at the discontinuous points. However, in many cases the viscosity solution is continuous and unique, especially in the interior domain. Therefore, it would be desirable to further prove the continuity of the solution (if this is true). The strong comparison result serves as a powerful tool to prove the continuity and uniqueness of the viscosity solution.

The strong comparison result allows us to compare any pair of viscosity subsolution and supersolution in a region inside the solution domain. The notation of strong comparison result is given as follows:

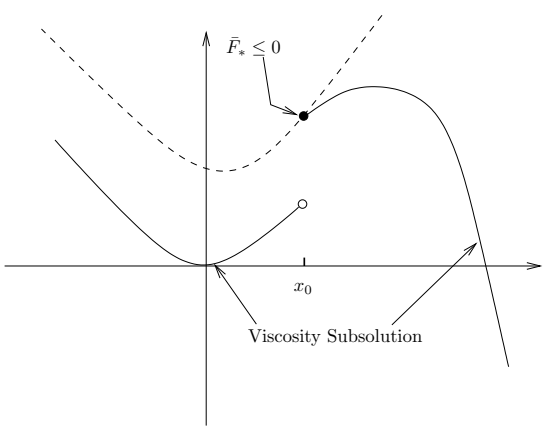
Definition 3.13 (Strong Comparison Result)**.** The strong comparison result for problem (3.1.16-3.1.17) holds in a domain $\Omega' \subseteq \bar{\Omega}$ if and only if for any usc subsolution f and any lsc supersolution g, we have

$$f \leq g \qquad \text{in } \Omega'. \tag{3.1.43}$$

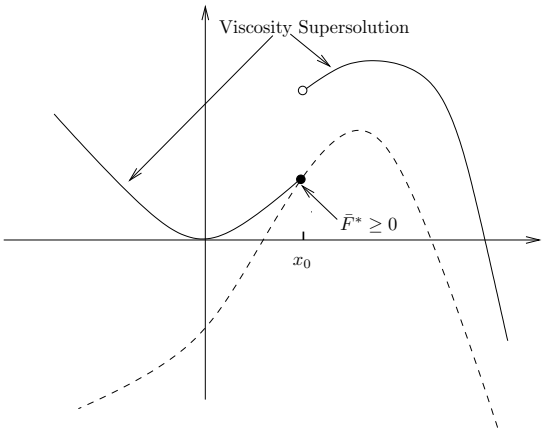
The strong comparison result immediately implies the following continuous and uniqueness result:



(A) A Viscosity Solution



(B) The Corresponding Viscosity Subsolution



(C) The Corresponding Viscosity Supersolution

FIGURE 3.4: *Illustration of the discontinuous viscosity solution definition. The curve in 3.4a represents a viscosity solution V that is discontinuous at x_0. The solid curve in 3.4b is the corresponding viscosity subsolution, which is the usc envelope V^* of the solution in 3.4a. The dash curve in 3.4b represents a smooth test function ϕ that touches the viscosity subsolution from above at the discontinuous point x_0, and then $\bar{F}_*\big(D^2\phi(x_0), D\phi(x_0), V^*(x_0), x_0\big) \leq 0$. The solid curve in 3.4c is the corresponding viscosity supersolution, which is the lsc envelope V_* of the solution in 3.4a. The dash curve in 3.4c represents a smooth test function χ that touches the viscosity supersolution from below at the discontinuous point x_0, and then $\bar{F}^*\big(D^2\chi(x_0), D\chi(x_0), V_*(x_0), x_0\big) \geq 0$.*

Theorem 3.14 (Continuity and Uniqueness of the Viscosity Solution)**.** *If the strong comparison result, as defined in Definition 3.13, for problem (3.1.16-3.1.17) holds in a domain $\Omega' \subseteq \bar{\Omega}$, then there exists a continuous and unique viscosity solution of (3.1.16-3.1.17) in Ω'.*

Proof. The existence of a viscosity solution is implied by Theorem 3.24. We will prove the continuity and uniqueness below.

Suppose V is a viscosity solution. We show V is continuous in Ω'. By definition, V^* and V_* are the viscosity subsolution and supersolution, respectively. Therefore, the strong comparison result given in Definition 3.13 implies that

$$V^* \leq V_* \qquad \text{in } \Omega'. \tag{3.1.44}$$

However, according to Remark 3.8, we obtain

$$V_* \leq V^*. \tag{3.1.45}$$

As a result, we obtain $V^* = V_* = V$ in Ω', which means V is continuous in Ω'.

Now assume there is another viscosity solution U. Following the above argument we can also obtain $U^* = U_* = U$ in Ω'. The strong comparison result and the continuity of solution implies that

$$U = U_* \geq V^* = V \qquad \text{in } \Omega' \tag{3.1.46}$$

and

$$U = U^* \leq V_* = V \qquad \text{in } \Omega'. \tag{3.1.47}$$

This implies $U = V$ in Ω'. Therefore, the viscosity solution is also unique in Ω'.

□

Next we discuss the existence of the strong comparison result for the gas storage problem (3.1.16-3.1.17) (or PDE (2.2.12) together with boundary equations (2.2.13-2.2.17)). There are various research papers deriving a strong comparison result for second-order

HJB equations associated with several types of boundary conditions [8, 10, 21, 7, 15]. In particular, [8, 10] prove that the viscosity solution of degenerate elliptic HJB equations with Dirichlet boundary conditions satisfies the strong comparison result, provided that several assumptions are satisfied. In [10], the author demonstrates that

S1 if the coefficient of the diffusion term (in our case $(\hat{\sigma}(P)P)^2$) vanishes at a region on a boundary with an outgoing or zero characteristic, independent of the value for the control variable, then the viscosity solution on this boundary region is the limit of the viscosity solution from interior points;

S2 if the characteristic at a region on the boundary, associated with the first order term in the PDE, is incoming to the domain independent of the choice of the control value, then the viscosity solution at the region corresponds to the specified boundary data in the classical sense.

We can regard the two-dimensional parabolic PDE (2.2.12) as a three-dimensional degenerate elliptic PDE in the variable $\mathbf{x} = (P, I, \tau) \in [0, P_{\max}] \times [0, I_{\max}] \times [0, T]$. The resulting elliptic PDE is degenerate in the sense that the equation does not contain the second-order derivatives with respect to τ and I, or, equivalently, the effective volatility (i.e. the diffusion term) is zero with respect to τ and I. We solve PDE (2.2.12) in the boundary region $\mathbf{x} \in [0, P_{\max}] \times \{0, I_{\max}\} \times (0, T)$. Conditions (2.2.14-2.2.15) imply that the statement **S1** above is satisfied for this boundary region. In the boundary region $\mathbf{x} \in \{0, P_{\max}\} \times [0, I_{\max}] \times (0, T)$ we solve equation (2.2.16-2.2.17). Since $\alpha(K(t) - P) \geq 0$ as $P \to 0$ and $\alpha(K(t) - P) \leq 0$ as $P \to P_{\max}$, then statement **S1** above is also satisfied for this region. Thus, the viscosity solution does not require boundary data in both P and I directions, which confirms our intuition in setting the boundary conditions in these directions. PDE (2.2.12) implies that statement **S2** above is satisfied in the region when $\mathbf{x} \in [0, P_{\max}] \times [0, I_{\max}] \times \{0\}$. This means that the viscosity solution uses the Dirichlet boundary condition, which we provided as the payoff function in equation (2.2.13).

From the analysis above, the boundary conditions we apply for equation (2.2.12) are in accordance with the behaviour of the viscosity solution at the boundary. Consequently,

we can use the strong comparison result in [8, 10] if equations (2.2.12-2.2.17) satisfy assumptions given in [8, 10]. However, a technical difficulty arises when we try to verify an assumption among those outlined in [8, 10]: the boundary is assumed to be smooth in [8, 10] so that the distance function from a point in the interior to the boundary is well-defined. In our case, however, the boundary surface is a cuboid, which results in non-smoothness of the distance function in the corners of the cuboid. In [21], the strong comparison result is proved for a similar (but not identical) problem associated with a non-smooth boundary. Consequently, we make the following assumption which is necessary to ensure that a unique viscosity solution to equation (2.2.12) exists.

Assumption 3.15. The gas storage pricing problem (3.1.16-3.1.17) (or pricing equation (2.2.12) and the associated boundary conditions (2.2.13-2.2.17)) satisfy the strong comparison result, as defined in Definition 3.13, in domain $\bar{\Omega} = [0, P_{\max}] \times [0, I_{\max}] \times [0, T]$.

3.2 Convergence to the Viscosity Solution

After presenting the definition of viscosity solutions in the previous section, we need to discuss the convergence of our scheme to the viscosity solution of the pricing problem.

Provided a strong comparison result for the PDE applies, the authors of [11, 6] demonstrate that a numerical scheme will converge to the viscosity solution of the equation if it is l_∞-stable, consistent, and monotone. Schemes failing to satisfy these conditions may converge to non-viscosity solutions. In fact, [75] gives an example where seemingly reasonable discretizations of nonlinear option pricing PDEs that do not satisfy the sufficient convergence conditions for viscosity solutions either never converge or converge to a non-viscosity solution. In this section, we review the notation of l_∞-stability, consistency and monotonicity from [11, 6] and verify that our fully implicit semi-Lagrangian scheme satisfies these properties.

As explained in Section 2.3.3, higher than or equal to third-order (quadratic) interpolation is needed for the operation Φ^{n+1} in (2.3.23) in order to achieve a second-order global truncation error for Crank-Nicolson timestepping (for smooth solutions). This

makes this scheme non-monotone in general, and hence we cannot guarantee convergence of high-order Crank-Nicolson timestepping to the viscosity solution because monotonicity can be obtained only for linear interpolation. We will, nevertheless, prove the consistency of the Crank-Nicolson timestepping scheme and carry out numerical experiments with Crank-Nicolson timestepping using quadratic interpolation.

We can write the discrete equations (2.3.25) at each node (P_i, I_j, τ^{n+1}), $n+1 \geq 1$, in a uniform format as

$$\begin{aligned}
&\mathcal{G}_{i,j}^{n+1}\big(h, V_{i,j}^{n+1}, \{V_{k,j}^{n+1}\}_{k\neq i}, \{V_{i,j}^{n}\}\big) \\
&\equiv \inf_{(\zeta_{i,j}^{n}, \zeta_{i,j}^{n+1}) \in C_j^{n+1}} \left\{ \frac{V_{i,j}^{n+1} - \big[\Phi^{n+1} V^{n}\big]_{i,j}}{\Delta\tau} - (1-\theta)\big[L^{n+1}V^{n+1}\big]_{i,j} - \theta\big[\Phi^{n+1}L^{n}V^{n}\big]_{i,j} \right. \\
&\qquad \left. - (1-\theta)\big[\big(\zeta_j^{n+1} - a\big(\zeta_j^{n+1}\big)\big)P\big]_i - \theta\big[\big(\zeta_j^{n} - a\big(\zeta_j^{n}\big)\big)P\big]_i \right\} \\
&= 0 \qquad\qquad \text{if } n+1 \geq 1,
\end{aligned} \tag{3.2.1}$$

where $\{V_{k,j}^{n+1}\}_{k\neq i}$ is the set of values $V_{k,j}^{n+1}$, $k \neq i$, $k = 0, \ldots, i_{\max}$ and $\{V_{i,j}^{n}\}$ is the set of values $V_{i,j}^{n}$, $i = 0, \ldots, i_{\max}$, $j = 0, \ldots, j_{\max}$. We also define $\mathcal{G}_{i,j}^{n+1}$ at payoff time $\tau = 0$ as

$$\mathcal{G}_{i,j}^{n+1}\big(h, V_{i,j}^{n+1}, \{V_{k,j}^{n+1}\}_{k\neq i}, \{V_{i,j}^{n}\}\big) \equiv V_{i,j}^{n+1} - B(P_i, I_j) = 0, \quad \text{if } n+1 = 0, \tag{3.2.2}$$

where $B(P_i, I_j)$ is the value of payoff at a node (P_i, I_j). Consequently, $\mathcal{G}_{i,j}^{n+1}$ completely specifies our semi-Lagrangian discretization.

3.2.1 l_∞-Stability

Definition 3.16 (l_∞-Stability)**.** Discretization (3.2.1-3.2.2) is l_∞-stable if

$$\|V^{n+1}\|_\infty \;\leq\; C_4 \;, \tag{3.2.3}$$

for $0 \leq n \leq N-1$ as $\Delta\tau \to 0$, $\Delta P_{\min} \to 0$, $\Delta I_{\min} \to 0$, where C_4 is a constant independent of $\Delta\tau$, $\Delta P_{\min}$, $\Delta I_{\min}$. Here $\|V^{n+1}\|_\infty = \max_{i,j} |V_{i,j}^{n+1}|$.

The stability of the semi-Lagrangian fully implicit discretization (3.2.1-3.2.2) is a consequence of the following Lemma.

Lemma 3.17 (l_∞ Stability of the Fully Implicit Scheme)**.** *Assuming that discretization (2.3.4) satisfies the positive coefficient condition (2.3.5) and linear interpolation is used in operation Φ^{n+1} in (2.3.23), then the scheme (3.2.1-3.2.2) satisfies*

$$\|V^{n+1}\|_\infty \leq \|V^0\|_\infty + C_5 \tag{3.2.4}$$

in the case of fully implicit timestepping ($\theta = 0$), where

$$C_5 = T \cdot P_{\max} \cdot \max\{\left|c_{\max}(I_{\max})\right|, \left|c_{\min}(0)\right|\}.$$

Proof. The proof directly follows from applying the maximum principle to the discrete equation (2.3.25). We omit the details here. Readers can refer to [40, Theorem 5.5] and [46] for complete stability proof of the semi-Lagrangian fully implicit scheme for American Asian options and that of finite difference schemes for controlled HJB equations, respectively. □

3.2.2 Consistency

Following [11, 6], we give a definition for the consistency of a discretization.

Definition 3.18 (Consistency)**.** The scheme $\mathcal{G}_{i,j}^{n+1}\big(h, V_{i,j}^{n+1}, \{V_{k,j}^{n+1}\}_{k\neq i}, \{V_{i,j}^n\}\big)$ given in equation (3.2.1-3.2.2) is consistent if, for all $\hat{\mathbf{x}} = (\hat{P}, \hat{I}, \hat{\tau}) \in \bar{\Omega} = [0, P_{\max}] \times [0, I_{\max}] \times [0, T]$ and any function $\phi(P, I, \tau)$ having bounded derivatives of all orders in $(P, I, \tau) \in \bar{\Omega}$ with $\phi_{i,j}^{n+1} = \phi(P_i, I_j, \tau^{n+1})$ and $\mathbf{x} = (P_i, I_j, \tau^{n+1})$, we have

$$\limsup_{\substack{h\to 0\\ \mathbf{x}\to\hat{\mathbf{x}}\\ \xi\to 0}} \mathcal{G}_{i,j}^{n+1}\big(h, \phi_{i,j}^{n+1} + \xi, \{\phi_{k,j}^{n+1} + \xi\}_{k\neq i}, \{\phi_{i,j}^n + \xi\}\big) \leq \bar{F}^*\big(D^2\phi(\hat{\mathbf{x}}), D\phi(\hat{\mathbf{x}}), \phi(\hat{\mathbf{x}}), \hat{\mathbf{x}}\big), \tag{3.2.5}$$

and

$$\liminf_{\substack{h\to 0\\ \mathbf{x}\to\hat{\mathbf{x}}\\ \xi\to 0}} \mathcal{G}_{i,j}^{n+1}\big(h, \phi_{i,j}^{n+1}+\xi, \{\phi_{k,j}^{n+1}+\xi\}_{k\neq i}, \{\phi_{i,j}^{n}+\xi\}\big) \geq \bar{F}_*\big(D^2\phi(\hat{\mathbf{x}}), D\phi(\hat{\mathbf{x}}), \phi(\hat{\mathbf{x}}), \hat{\mathbf{x}}\big), \quad (3.2.6)$$

where $\bar{F}$ is defined in (3.1.39) and $\bar{F}^*$ and $\bar{F}_*$ are respectively the usc and lsc envelopes of $\bar{F}$, as defined in Definition 3.7.

The consistency of scheme (3.2.1-3.2.2) is given in the following Lemma:

Lemma 3.19 (Consistency)**.** *Suppose the mesh size and timestep size satisfy equations (2.3.1), and the control parameters satisfy condition (2.2.7). Then the discretization (3.2.1-3.2.2) is consistent as defined in Definition 3.18, provided that $c_{\max}$, $c_{\min}$ and $a(c)$ satisfy equations (2.2.1), (2.2.3) and (2.2.5), respectively. In particular, assuming the solution is smooth and that linear interpolation is used in operation Φ^{n+1} in (2.3.23), the global discretization error of the scheme $\mathcal{G}_{i,j}^{n+1}$ is $O(h)$.*

Proof. See Appendix D. □

Remark 3.20 (Consistency of the Bang-Bang and No Bang-Bang Methods)**.** *As introduced in Section 2.4, we solve the local optimization problem in scheme (3.2.1) using the bang-bang or no bang-bang approach. This introduces additional numerical errors.*

According to Remark 2.4, we can still verify Definition 3.18 for the no bang-bang method since it solves the optimization problem consistently.

On the other hand, the bang-bang method consistently solves the local optimization problem corresponding to the equation (2.4.9). Consequently, we can also verify the consistency definition for equation (2.4.9). Using the results in Theorem 3.25, we can prove the bang-bang method with a fully implicit timestepping converges to the viscosity solution of (2.4.9). Since equation (2.4.9) has the same viscosity solution as equation (2.2.12), the bang-bang method also converges to the viscosity solution of PDE (2.2.12).

3.2.3 Monotonicity

In this section, we discuss the monotonicity of the fully implicit scheme

Definition 3.21 (Monotonicity)**.** The discretization $\mathcal{G}_{i,j}^{n+1}\big(h, V_{i,j}^{n+1}, \{V_{k,j}^{n+1}\}_{k\neq i}, \{V_{i,j}^{n}\}\big)$ given in equation (3.2.1-3.2.2) is monotone if

$$\begin{aligned}&\mathcal{G}_{i,j}^{n+1}\big(h, V_{i,j}^{n+1}, \{X_{k,j}^{n+1}\}_{k\neq i}, \{X_{i,j}^{n}\}\big)\\ &\leq \mathcal{G}_{i,j}^{n+1}\big(h, V_{i,j}^{n+1}, \{Y_{k,j}^{n+1}\}_{k\neq i}, \{Y_{i,j}^{n}\}\big);\quad \text{for all } X_{i,j}^{n} \geq Y_{i,j}^{n},\ \forall i,j,n.\end{aligned} \tag{3.2.7}$$

This definition of monotonicity is equivalent to that introduced in [11, 6].

Lemma 3.22 (Monotonicity)**.** *If the discretization (2.3.4) satisfies the positive coefficient condition (2.3.5) and linear interpolation is used in operation Φ^{n+1} in (2.3.23), then in the case of fully implicit timestepping ($\theta = 0$), the discretization $\mathcal{G}_{i,j}^{n+1}\big(h, V_{i,j}^{n+1}, \{V_{k,j}^{n+1}\}_{k\neq i}, \{V_{i,j}^{n}\}\big)$, as given in (3.2.1-3.2.2), is monotone according to Definition 3.21.*

Proof. The proof directly follows that of monotonicity of finite difference schemes for controlled HJB equations in [9, 46]. □

3.2.4 Arbitrage Inequalities

The authors of [25, 26] demonstrate that a financially meaningful discretization should satisfy arbitrage inequalities, which means the inequality of contract payoffs is preserved in the inequalities of contract values. Our scheme (3.2.1-3.2.2) in terms of fully implicit timestepping ($\theta = 0$) satisfies the following arbitrage inequalities.

Theorem 3.23 (Discrete Arbitrage Inequalities)**.** *If the discretization (2.3.4) satisfies the positive coefficient condition (2.3.5) and linear interpolation is used in operation Φ^{n+1} in (2.3.23), then in the case of fully implicit timestepping ($\theta = 0$), the discretization (3.2.1-3.2.2) satisfies a discrete comparison principle. That is, if $V^n > W^n$ and V^{n+1}, W^{n+1} satisfy (3.2.1-3.2.2), then $V^{n+1} > W^{n+1}$.*

Proof. The proof directly follows from the approach in [40, Theorem 6.2]. □

3.2.5 Convergence

Lemmas 3.17, 3.19 and 3.22 and the results in [11, 6] directly imply the following properties of the solution:

Theorem 3.24. *Assume that discretization (3.2.1-3.2.2) satisfies all the condition required for Lemmas 3.17, 3.19 and 3.22. Let*

$$\overline{V}(P,I,\tau) = \limsup_{\substack{h\to 0\\ P_i\to P\\ I_j\to I\\ \tau^{n+1}\to\tau}} V_{i,j}^{n+1} \qquad \underline{V}(P,I,\tau) = \liminf_{\substack{h\to 0\\ P_i\to P\\ I_j\to I\\ \tau^{n+1}\to\tau}} V_{i,j}^{n+1}. \tag{3.2.8}$$

Then $\overline{V}$ and $\underline{V}$ are respectively the viscosity subsolution and supersolution for the gas storage problem (3.1.16-3.1.17) (or PDE (2.2.12) together with boundary equations (2.2.13-2.2.17)) in the closed domain $(P,I,\tau) \in \bar{\Omega} = [0, P_{\max}] \times [0, I_{\max}] \times [0,T]$ in the case of fully implicit timestepping ($\theta = 0$).

Theorem 3.24 reveals that in theory the viscosity solution of the problem can be constructed from the numerical solution of our scheme. If the strong comparison result to the problem holds, then our scheme will converge to the unique and continuous viscosity solution, as stated by the following Theorem:

Theorem 3.25 (Convergence to the Viscosity Solution)**.** *If all conditions in Theorem 3.24 are satisfied, and, in addition, Assumption 3.15 holds, then scheme (3.2.1-3.2.2) converges to the continuous and unique viscosity solution of the gas storage problem (3.1.16-3.1.17) in domain $\bar{\Omega} = [0, P_{\max}] \times [0, I_{\max}] \times [0,T]$ in the case of fully implicit timestepping ($\theta = 0$). In other words, we have*

$$\overline{V}(P,I,\tau) = \underline{V}(P,I,\tau) = V(P,I,\tau), \quad \forall (P,I,\tau) \in \bar{\Omega}. \tag{3.2.9}$$

Remark 3.26. *If the strong comparison result does not hold at some point $\mathbf{x} = (P,I,\tau) \in \bar{\Omega}$, then we have $\underline{V} < \overline{V}$ at $\mathbf{x}$ according to (3.2.8) and Definition 3.13. Hence our numerical solution does not converge at $\mathbf{x}$. However, this is expected because the viscosity solution itself is discontinuous and non-unique at $\mathbf{x}$. Consequently, Remark 3.12 shows that any*

value residing between $\underline{V}(\mathbf{x})$ and $\overline{V}(\mathbf{x})$ is a valid viscosity solution. Therefore, it is not clear which value the scheme should converge to at the point $\mathbf{x}$. In this case, the most precise information of the solution is the bounds $\underline{V}(\mathbf{x})$ and $\overline{V}(\mathbf{x})$, which can still be obtained (in theory) from our numerical solution.

In addition, the discontinuity of the viscosity solution normally occurs only on the boundaries when the boundary equations are degenerate elliptic. As explained in [25], the convergence of the numerical scheme at these points often has no practical importance.

3.3 Summary

Our work in this chapter is summarized as follows:

- We introduce the notation of (possibly discontinuous) viscosity solutions that is able to handle various types of boundary conditions.
- We prove that the fully implicit, semi-Lagrangian scheme is unconditionally l_∞-stable, monotone and consistent. Therefore, provided a strong comparison property holds, the fully implicit, semi-Lagrangian discretization converges to the unique and continuous viscosity solution of the pricing equation using the results in [11, 6].

Chapter 4

Numerical Results for the Gas Storage Valuation Problem

Having presented several semi-Lagrangian discretization schemes in the previous chapter, in this chapter we conduct numerical experiments based on these schemes.

We use "dollars per million British thermal unit" ($/mmBtu) and "million cubic feet" (MMcf) as the default units for gas spot price P and gas inventory I, respectively. Since 1000 mmBtus are roughly equal to 1 MMcf, in order to unify the units, we need to multiply gas spot price by 1000 when computing payoffs or revenues.

Throughout the numerical experiments, we use the following non-smooth payoff function from [18]

$$V(P, I, t = T) = -2P \max(1000 - I, 0). \tag{4.0.1}$$

Equation (4.0.1) indicates that severe penalties are charged if the gas inventory is less than 1000 MMcf and no compensation is received when the inventory is above 1000 MMcf. Naturally, such a payoff structure will force the operator of a gas storage facility to maintain the gas inventory as close to 1000 MMcf as possible at maturity to avoid revenue loss.

This chapter is arranged as follows: we first give numerical results for the case without incorporating the seasonality effect into the equilibrium natural gas spot price; we then

incorporate the seasonality feature and illustrate its influence on both the solution value and the optimal control strategy. At the end of this chapter, we further extend the underlying risk neutral gas spot price process to include a compound Poisson process that simulates random jumps of the gas prices, and then present numerical results incorporating the jump diffusion process.

4.1 No Seasonality Effect

In this section, we assume that the equilibrium gas price is independent of time, that is, we set $K(t) = K_0$ in equation (2.2.10).

We first carry out a convergence analysis assuming that the risk neutral natural gas spot price follows the mean-reverting process (2.2.9) with $\alpha = 2.38$, $K_0 = 6$, $\sigma = 0.59$. In other words, the risk neutral gas spot price follows

$$dP = 2.38(6 - P)dt + 0.59PdZ. \tag{4.1.1}$$

We are most interested in the solution when the gas spot price is near the long-term equilibrium price, i.e., $P = 6$ \$/mmBtu for (4.1.1). Note that when $I = 1000$ MMcf, the payoff is non-smooth (see equation (4.0.1)). Consequently, to fully test our semi-Lagrangian discretization schemes, we focus on the convergence results at $(P, I) = (6, 1000)$. We use an unequally spaced grid in the P, I directions, where there are more nodes around the mesh point $(P, I) = (6, 1000)$, compared with other locations.

Table 4.1 lists other input parameters for pricing the value of a gas storage contract. The convergence results obtained from refining the mesh spacing and timestep size are shown in Table 4.2, where we use fully implicit and Crank-Nicolson timestepping schemes associated with both the bang-bang and no bang-bang methods for solving the discrete optimization problem in Algorithm 2.1. Linear interpolation and quadratic interpolation are used for fully implicit and Crank-Nicolson timestepping, respectively. (Refer to Section 2.3 for a discussion on interpolation schemes.) Following [75], in order to improve the convergence for non-smooth payoff (4.0.1), we use a modification suggested by [76, 49]

for Crank-Nicolson timestepping. Specifically, we apply fully implicit timestepping in the first four timesteps, and use Crank-Nicolson timestepping in the rest of the timesteps.

Parameter	Value	Parameter	Value
r	0.1	k_2	730000
T	3 years	k_3	500
$I_{\max}$	2000 MMcf	k_4	2500
k_1	2040.41	k_5	$1.7 \cdot 365$

TABLE 4.1: *Input parameters used to price the value of a gas storage contract, where $I_{\max}$ is the maximum storage inventory; k_1, k_2, k_3, k_4, k_5 are parameters in equations (2.2.1-2.2.2) and (2.2.5). The values of $I_{\max}, k_1, k_2, k_3, k_4, k_5$ are taken from [80].*

The results in Table 4.2 indicate that both timestepping schemes converge to the same solution, although convergence to the viscosity solution can only be guaranteed for fully implicit timestepping given Assumption 3.15. We define the convergence ratio as the ratio of successive changes in the solution, as the timestep and mesh size are reduced by a factor of two. A ratio of two indicates first-order convergence, while a ratio of four indicates second order convergence. The convergence ratios are approximately two for fully implicit timestepping with both the bang-bang and the no bang-bang methods. Note that the no bang-bang method is a more general approach which can be used in cases where controls are not of bang-bang type.

It is interesting to note that in a fixed refinement level, the bang-bang method results in a smaller value than the no bang-bang method for both timestepping schemes. This is because the no bang-bang method actually solves the discrete optimization problem in Algorithm 2.1, instead of only testing a finite set of points, which results in a higher solution value for PDE (2.2.12) (for a finite grid size) than the bang-bang method.

For the fully implicit tests in Table 4.2, the no bang-bang method requires about 10% more CPU time compared to the bang-bang method. This is consistent with our earlier estimates, since the the no bang-bang examines only a constant number of grid nodes per node in order to solve the local optimization problem.

Table 4.2 also shows that Crank-Nicolson timestepping does not appear to converge at a second-order rate. We have observed this same effect in many of our tests. Since

P grid nodes	I grid nodes	No. of timesteps	Bang-bang method		No bang-bang method	
			Value	Ratio	Value	Ratio
Fully implicit timestepping						
53	61	500	4477036	n.a.	4556380	n.a.
105	121	1000	4503705	n.a.	4542845	n.a.
209	241	2000	4514723	2.42	4534660	1.63
417	481	4000	4519809	2.17	4530331	1.89
833	961	8000	4522653	1.79	4528219	2.05
Crank-Nicolson timestepping						
53	61	500	4483667	n.a.	4520475	n.a.
105	121	1000	4509076	n.a.	4525352	n.a.
209	241	2000	4517960	2.86	4526280	5.26
417	481	4000	4522632	1.90	4527225	0.98
833	961	8000	4524948	2.02	4527331	8.92

TABLE 4.2: *The value of a natural gas storage facility at $P = 6$ \$/mmBtu and $I =$ 1000 MMcf. The risk neutral gas spot price follows the mean-reverting process (4.1.1). Convergence ratios are presented for fully implicit and Crank-Nicolson timestepping schemes with the bang-bang and the no bang-bang methods. Constant timesteps are used. The payoff function is given in (4.0.1). Other input parameters are given in Table 4.1. Crank-Nicolson incorporates the modification suggested in [76].*

we do not seem to obtain any benefit from Crank-Nicolson timestepping, fully implicit timestepping appears to be a better choice since we are guaranteed convergence to the viscosity solution given Assumption 3.15, as shown in Section 3.2. In the rest of this thesis, we will use fully implicit timestepping exclusively.

Figure 4.1 shows the optimal control surface at $t = 0$ as a function of P and I. This surface is similar to that given in [80]. The interpretation given in [80] also applies to Figure 4.1.

Our numerical computations truncate the domain $P \in [0, \infty]$ to $[0, P_{\max}]$. In order to test the influence of the domain truncation on the solution, we compute the solution values at $P = 6$ \$/mmBtu, $I = 1000$ MMcf using two different values of $P_{\max}$: $P_{\max} = 2000$ and 20000 \$/mmBtu. We found that for all four refinement levels, the first ten digits of the two solution values are identical. This indicates that by setting $P_{\max} = 2000$ \$/mmBtus, there is a negligible solution error incurred by the domain truncation. As a result, all subsequent results will be reported using $P_{\max} = 2000$ \$/mmBtus.

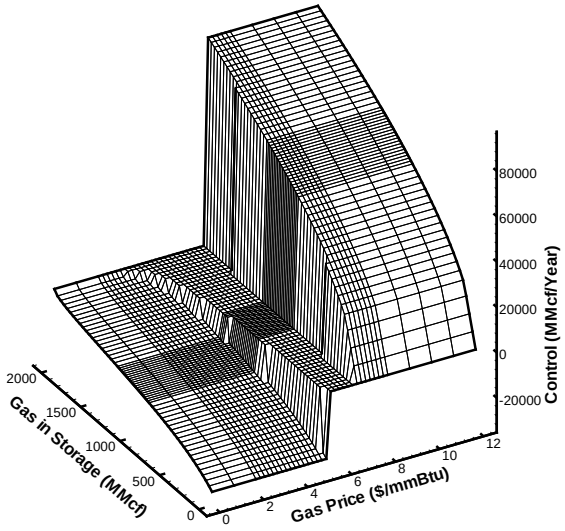


FIGURE 4.1: *The optimal control strategy at current time $t = 0$ as a function of gas spot price P and gas inventory I. The risk neutral gas spot price follows the mean-reverting process (4.1.1). Payoff function is given in (4.0.1). Other input parameters are given in Table 4.1. Fully implicit timestepping with the no bang-bang method and with constant timesteps is used.*

If an explicit semi-Lagrangian scheme is used, then the stability condition is

$$\Delta\tau^n < \min_i \left\{ \frac{1}{\gamma_i^n + \beta_i^n + r} \right\}, \tag{4.1.2}$$

where $\Delta\tau^n = \tau^{n+1} - \tau^n$ and parameters γ_i^n, β_i^n are given in Appendix B. Condition (4.1.2) implies that $\Delta\tau = O((\Delta P_{\min})^2)$, where $\Delta P_{\min} = \min_i(P_{i+1} - P_i)$. In contrast, there is no such timestep restriction for fully implicit timestepping.

In [80], a fully explicit method (standard timestepping) is used for the gas storage problem. In this case, the stability condition would be $\Delta\tau = O((\Delta P_{\min})^2 + \Delta I_{\min})$, where $\Delta I_{\min} = \min_j(I_{j+1} - I_j)$.

Recall from Remark 2.6 that both the implicit method used here and an explicit method have complexity linear in the number of space nodes per timestep. In general, the constant in the complexity estimate will favour the explicit, standard timestepping scheme [80], due to the the extra interpolation operations required by the implicit method.

In terms of running time, we expect that the implicit method will be superior to the explicit method if the spatial error dominates, since the explicit stability condition will force smaller timesteps than is required for accuracy.

On the other hand, there will undoubtedly also be cases where the error is dominated by the timestepping error, in which case an explicit method may require less running time.

We remark that the fully implicit method has the practical advantage that we are completely free to place (P, I) nodes wherever is deemed necessary, since this has no effect on the permitted timestep size.

4.2 Incorporating the Seasonality Effect

In this section, we present numerical results after incorporating the seasonality effect into the equilibrium price of the mean-reverting process (4.1.1). We modified process (4.1.1) to

$$dP = 2.38(6 + \sin(4\pi t) - P)dt + 0.59PdZ, \tag{4.2.1}$$

where the additional term $\sin(4\pi t)$ makes the equilibrium price a periodic function to represent summer and winter peaks in the equilibrium price. The convergence results for this case are shown in Table 4.3. Comparing Table 4.3 with Table 4.2 indicates that incorporating the seasonality component does not affect the convergence ratio, but does increase the solution value for a fixed refinement level. This is reasonable, since the seasonality effect gives the operator of a gas storage facility an opportunity for obtaining greater profits by using an optimal strategy that takes advantage of the seasonality feature. For example, a simple strategy of buying and storing gas in spring and then producing and selling gas in summer can normally produce profits from the seasonality effect.

Figure 4.2 shows the optimal control strategy in the seasonality case that evolves over time as a function of P when the inventory is fixed at $I = 1000$ MMcf. The figure suggests that the optimal strategy is to inject gas at the maximum rate (corresponding to the negative control region in the surface) when the gas price is low, to produce gas at the maximum rate (corresponding to the positive control region) when the gas price

P grid nodes	I grid nodes	No. of timesteps	Bang-bang method		No bang-bang method	
			Value	Ratio	Value	Ratio
53	61	500	4815891	n.a.	4889602	n.a.
105	121	1000	4839796	n.a.	4875449	n.a.
209	241	2000	4848843	2.64	4866953	1.67
417	481	4000	4853100	2.13	4862530	1.92
833	961	8000	4855485	1.78	4860397	2.07

TABLE 4.3: *The value of a natural gas storage facility at $P = 6$ \$/mmBtu and $I = 1000$ MMcf. The risk neutral gas spot price follows the mean-reverting process (4.2.1) that incorporates the seasonality effect. Convergence ratios are presented for fully implicit timestepping with the bang-bang and the no bang-bang methods. Constant timesteps are used. The payoff function is given in (4.0.1). Other input parameters are given in Table 4.1.*

is high, and to do nothing (corresponding to the zero control region) when the gas price is near the long-term equilibrium price. From the figure, we can clearly notice the effect of the seasonality on the control strategy: the boundary curve between the zero control region and the negative/positive control region, which represents the control switching boundary between no operation and injecting/producing gas, is periodic when it is far from maturity. We can also observe that when the contract is close to maturity, the zero control region expands rapidly. This phenomenon is caused by the payoff function (4.0.1): at $I = 1000$ MMcf, when close to maturity, the operator tends to stop producing gas to avoid the severe penalty at maturity. In addition, the operator will stop injecting, since any leftover gas is lost to the operator.

To illustrate the difference of the optimal control strategies before and after incorporating the seasonality effect, Figure 4.3 shows the control switching boundary curves at $I = 1000$ MMcf as a function of time to maturity with respect to processes (4.1.1) and (4.2.1), respectively.

4.3 Incorporating the Jump Effect

It is not uncommon to see spot gas price jumps, when gas is used to power electrical generating plants in times of high electricity demand. Spot gas price can jump by as much

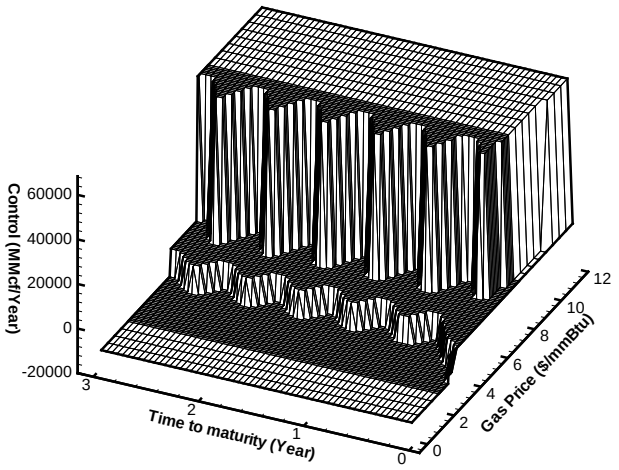


FIGURE 4.2: *The Optimal control strategy as a function of time to maturity $\tau = T - t$ and gas spot price P when the gas inventory resides at $I = 1000$ MMcf. The risk neutral gas spot price follows the mean-reverting process (4.2.1), with seasonality. The payoff function is given in (4.0.1). Other input parameters are given in Table 4.1. Fully implicit timestepping with the no bang-bang method and with constant timesteps is used.*

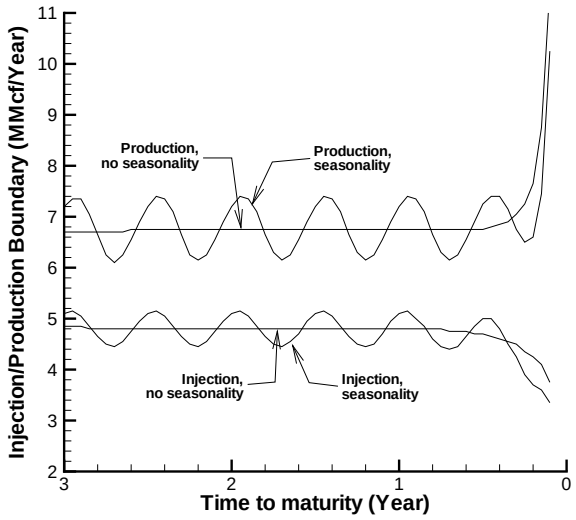


FIGURE 4.3: *Control switching boundary curves as a function of time to maturity $\tau = T - t$ with respect to processes (4.1.1) (without incorporating the seasonality effect) and (4.2.1) (incorporating the seasonality effect) at $I = 1000$ MMcf. The payoff function is given in (4.0.1). Other input parameters are given in Table 4.1. Fully implicit timestepping with the no bang-bang method and with constant timesteps is used.*

as 20% in a single day. To model this effect, in this section, we take the mean-reverting process (2.2.9) and extend it to include a compound Poisson process representing the jump effect, and present numerical results including a jump diffusion process. After adding a jump component, process (2.2.9) becomes

$$dP = [\alpha(K(t) - P) - \lambda\kappa P]dt + \sigma P dZ + (\eta - 1)P dq, \tag{4.3.1}$$

where

- dq is the independent Poisson process $= \begin{cases} 0 & \text{with probability } 1 - \lambda dt, \\ 1 & \text{with probability } \lambda dt, \end{cases}$
- λ is the jump intensity representing the mean arrival time of the Poisson process,
- η is a random variable representing the jump size of gas price—when $dq = 1$, price jumps from P to $P\eta$. We assume that η follows a probability density function $g(\eta)$,
- κ is $E[\eta - 1]$, where $E[\,\cdot\,]$ is the expectation operator.

Assuming that the risk neutral gas spot price follows the jump diffusion process (4.3.1), the pricing PDE (2.2.12) turns into the following controlled partial integro-differential equation (PIDE)

$$\begin{aligned} V_\tau = \frac{1}{2}\sigma^2 P^2 V_{PP} + [\alpha(K(t) - P) - \lambda\kappa P]V_P + \sup_{c \in C(I)} \{(c - a(c))P - (c + a(c))V_I\} \\ - rV + \left(\lambda \int_0^\infty V(P\eta)g(\eta)d\eta - \lambda V\right). \end{aligned} \tag{4.3.2}$$

Since there is no control variable in the integral term of PIDE (4.3.2), we can use the methods described in [40, 41, 39] to extend the semi-Lagrangian discretization schemes introduced in Chapter 2 to solve the PIDE without difficulty. We note that it is straightforward to combine the methods in Chapter 3 with the approaches in [40] to show that the resulting scheme is consistent, stable and monotone.

During our numerical experiments, we assume that the probability density function $g(\eta)$ follows a log-normal distribution

$$g(\eta) = \frac{1}{\sqrt{2\pi}\gamma\eta} \exp\Big(-\frac{(\log(\eta) - \nu)^2}{2\gamma^2}\Big) \tag{4.3.3}$$

with expectation $E[\eta] = \exp(\nu + \gamma^2/2)$. We will choose values of the parameters ν and γ such that $(\eta - 1)$, the relative change in the gas spot price, has mean zero and variance 0.04.

Table 4.4 lists values of the parameters for process (4.3.1) and for the log-normal density function (4.3.3), where the parameters of the drift and diffusion components in process (4.3.1) take the same values as those in process (4.2.1) for the case without incorporating the jump effect. Note that we set the jump intensity $\lambda = 12$ so that random jumps appear approximately once every month. Table 4.5 presents the convergence results for the solution to PIDE (4.3.2). Table 4.5 shows that the jump effect greatly increases the value of the storage facility.

Since the controls in equation (4.3.2) do not appear in the integral terms, it seems reasonable to suppose that the converged controls for problem (4.3.2) will also be of the bang-bang type, but we are not aware of a proof of this. We will solve (4.3.2) using both bang-bang and no bang-bang methods, and our numerical results verify that both techniques converge to the same solution.

For the finest grid in Table 4.2, the no bang-bang method with jumps takes about three times more CPU time compared to the same problem with no jumps. This is simply because we need several iterations per timestep to solve the fully implicit discretized equations, including the jump term [40]. Each iteration requires one tridiagonal linear system solve and two FFTs. Note that we can also evaluate the jump term explicitly to avoid the iterations [40]. The resulting scheme is still unconditionally stable, monotone and consistent, but it is first-order correct in time.

The results in Table 4.5 also indicate that the no bang-bang method achieves first-order convergence, but not the bang-bang method. To further study this behaviour, in

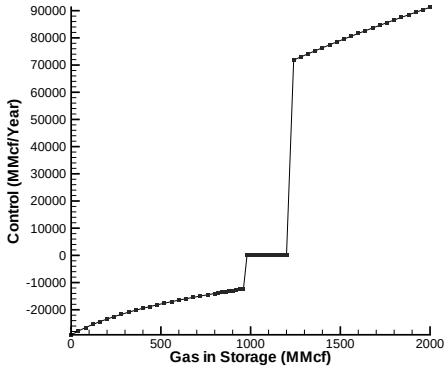
Figure 4.4, we show the control curves at $P = 6$ \$/mmBtu, $\tau = \tau^1 = .006$ year as a function of I, obtained using these two methods. We present the control curves produced using one timestep (i.e., $\Delta\tau = \tau^1$) with a coarse space grid, as well as a fine grid solution with many timesteps. From the figure, we can observe that when using finer grids (and more timesteps), both the bang-bang and the no bang-bang methods converge to the same control strategy. In contrast, when a coarse grid and one timestep are used, the control curves produced by both methods differ from the converged controls near $I = 1000$ MMcf (excluding $I = 1000$ MMcf), hence are not accurate. However, by actually solving the discrete optimization problem in Algorithm 2.1, the no bang-bang method produces a much smoother control curve on coarse grids, compared with the bang-bang method. Consequently, this would appear to explain why the no bang-bang method is able to generate a smoother solution for the value function as the grid size and timestep size is reduced compared to the bang-bang method.

Parameter	Value	Parameter	Value
α	2.38	ν	-0.0196
$K(t)$	$6 + \sin(4\pi t)$	γ	0.198
σ	0.59	λ	12

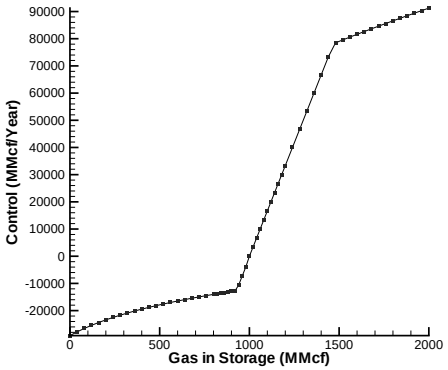
TABLE 4.4: *Input parameters for the jump diffusion process (4.3.1) and the log-normal density function (4.3.3). The parameters of the jump size density function are selected so that* $E[(\eta - 1)] = 0$ *and* $E[(\eta - 1)^2] = .04$.

P grid nodes	I grid nodes	No. of timesteps	Bang-bang method		No-bang-bang method	
			Value	Ratio	Value	Ratio
79	61	500	7995143	n.a.	8070698	n.a.
157	121	1000	7962386	n.a.	7999775	n.a.
313	241	2000	7951062	2.89	7971737	2.53
625	481	4000	7951032	377	7961554	2.75
1249	961	8000	7951976	-0.03	7957509	2.52

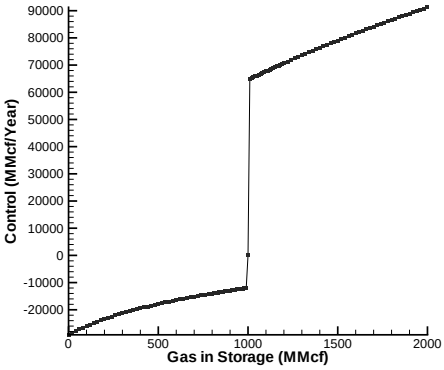
TABLE 4.5: *The value of a natural gas storage facility at* $P = 6$ *\$/mmBtu and* $I = 1000$ *MMcf. The risk neutral gas spot price follows the mean-reverting process (4.3.1) (incorporating the seasonality and the jump effects). Convergence ratios are presented for fully implicit timestepping with the bang-bang and the no bang-bang methods. Constant timesteps are used. The payoff function is given in (4.0.1). Input parameters are given in Tables 4.4 and 4.1.*



(A) Bang-bang method, $\tau = .006$, one timestep, coarse space grids



(B) No bang-bang method, $\tau = .006$, one timestep, coarse space grids



(C) Bang-bang or no bang-bang method, $\tau = .006$, fine space grid, and many timesteps.

FIGURE 4.4: *Control curves as a function of gas inventory I obtained at $\tau^1 = .006$ year with gas price $P = 6$ \$/mmBtu. The top panel shows the bang-bang and the no bang-bang methods with one timestep and a coarse space grid. The bottom panel shows the results for both bang-bang and no bang-bang methods, using a fine space grid and many timesteps. The risk neutral gas spot price follows the mean-reverting process (4.3.1). The payoff function is given in (4.0.1). Input parameters are given in Tables 4.4 and 4.1.*

Figure 4.5 compares the control switching boundary curves obtained before and after incorporating the jump effect when $I = 1000$ MMcf. The figure indicates that the zero control region (the region contained between two boundary curves) resulting from the jump diffusion process (4.3.1) is wider than that resulting from process (4.2.1). This occurs because, under the jump scenario, the operator is willing to wait for a jump in the gas price and then operate the facility after the jump to obtain more profit, which makes the zero control region wider. In addition, Figure 4.5 shows that the jump effect disappears when the contract is close to maturity because of the fear of revenue loss at maturity due to the payoff structure (4.0.1).

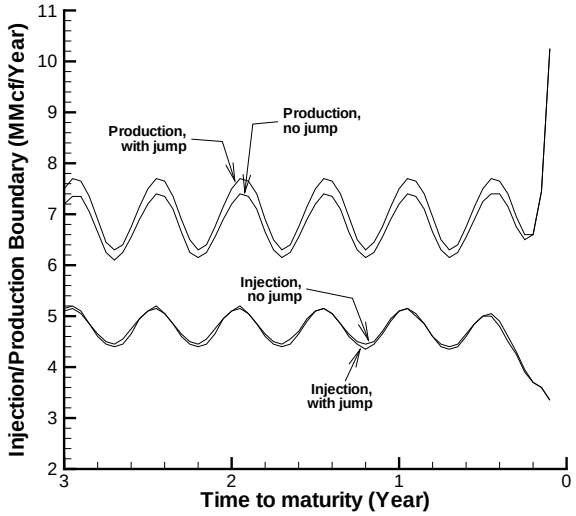


FIGURE 4.5: *Control switching boundary curves as a function of time to maturity $\tau = T - t$ with respect to processes (4.2.1) (without incorporating the jump effect) and (4.3.1) (incorporating the jump effect) at $I = 1000$ MMcf, where parameter values for process (4.3.1) are given in Table 4.4. Other input parameters are given in Table 4.1. Fully implicit timestepping with the no bang-bang method and with constant timesteps is used.*

4.4 Summary

Our contributions in this chapter are summarized as follows:

- We conduct numerical experiments based on our semi-Lagrangian discretizations

given in previous chapters. The numerical results indicate that fully implicit timestepping can achieve first-order convergence, while Crank-Nicolson timestepping does not appear to converge at a higher than first-order rate. Thus fully implicit timestepping is probably a better choice since it guarantees convergence to the viscosity solution and it is also straightforward to implement.

- We then extend the mean-reverting process (2.2.9) to include a compound Poisson process to model the jumps in gas prices. This results in a partial integrodifferential equation (PIDE) for gas storage valuation problem. Since semi-Lagrangian methods completely separate the inventory variable from the underlying stochastic component, we can easily incorporate the discretizations for the jump component, as described in [40, 41, 39], into the fully implicit semi-Lagrangian scheme to solve the pricing PIDE. The resulting scheme is still consistent, l_∞-stable and monotone.

Chapter 5

A Regime-Switching Model for Natural Gas Spot Prices

In Chapter 2 we value natural gas storage facilities assuming that natural gas spot prices follow a one-factor mean-reverting model. In this chapter, we propose a one-factor regime-switching model for natural gas spot and demonstrate by calibration that the regime-switching model is able to fit the market data more accurately than a typical one-factor mean-reverting model. In the next chapter we will solve the gas storage pricing problem under the regime-switching model.

Our primary objective in this chapter is to obtain reasonable parameters which will be used to carry out example computations for solving the gas storage HJB equation with a regime-switching model.

5.1 Introduction

Previous work on the valuation of natural gas storage facilities has almost exclusively assumed that natural gas spot prices follow one-factor mean-reverting processes. However, as demonstrated in [78, 57] and again this chapter, one-factor mean-reverting models do not seem to be able to capture the dynamics of typical gas forward curves. Consequently,

we need to resort to other more complex stochastic models for natural gas prices in order to more accurately price gas storage contracts.

A number of multi-factor models for the natural gas spot price are suggested in [91, 67, 57, 20]. More general multi-factor models for commodity spot prices are developed in [78, 27]. Nevertheless, it is computationally expensive to apply the two-factor and three-factor commodity spot price models in [78, 27] to price complex commodity derivatives such as the gas storage contracts, although they seem able to fit the market futures prices.

Consequently, we will focus on one-factor regime-switching models for natural gas spot prices. More precisely, we propose a model with a single Brownian motion process and a jump process. Initially proposed in [51], a regime-switching model consists of several regimes; within each regime the gas price follows a distinct stochastic process. The price process can randomly shift between these regimes due to various reasons, such as changes of weather conditions, alterations of demand and supply, and surprise events such as political instability. Regime-switching models have been used in several areas. For example, [52] develops a regime-switching model for equities. [50] and [5] use regime-switching processes to model term structures of interest rates. Various regime-switching models are calibrated to electricity spot prices in [37, 55, 36, 35, 77].

This chapter is arranged as follows: we first propose a one-factor mean-reverting model and a regime-switching model for natural gas spot prices. Then we calibrate the models to market futures data and examine the calibration performance. Finally, we obtain the values of model volatility by calibrating to market futures options.

5.2 Natural Gas Spot Price Models

In this section, we specify two one-factor models that we use to examine the dynamics of the natural gas spot price. Since we are interested in pricing derivatives on natural gas, we will consider directly the risk neutral price processes with parameters given under the $\mathbb{Q}$ measure.

5.2.1 One-Factor Mean-Reverting Model (MR Model)

Let P denote the natural gas spot price. In the MR model, the gas spot price follows a mean-reverting stochastic process with the seasonality effect represented in the drift term. The risk neutral gas spot price is modeled by a stochastic differential equation (SDE) given by

$$dP = \alpha(K_0 - P)dt + \sigma P dZ + S(t)Pdt \ , \tag{5.2.1}$$

$$S(t) = \beta_A \sin(2\pi(t - t_0 + C_A(t_0))) + \beta_{SA} \sin(4\pi(t - t_0 + C_{SA}(t_0))) \ , \tag{5.2.2}$$

where

- $\alpha > 0$ is the mean-reversion rate,
- $K_0 > 0$ is the long-term equilibrium price,
- $\sigma > 0$ is the volatility,
- dZ is an increment of the standard Gauss-Wiener process,
- $S(t)$ is a time-dependent term so that $S(t)Pdt$ is the price change at time t contributed by the seasonality effect. Note that multiplying $S(t)$ with P guarantees the price of natural gas always stays positive. This is a useful property during our calibration process,
- β_A is the annual seasonality parameter,
- t_0 is a reference time satisfying $t_0 < t$.
- $C_A(t_0)$ is the annual seasonality centering parameter for t_0. We define

$$C_A(t_0) = A_0 + D(t_0), \tag{5.2.3}$$

  where A_0 is a constant time adjustment parameter obtained through calibration; $D(t_0)$ is the distance between the reference time t_0 and the first date in January in

the year of t_0. Thus, by calibrating the value of A_0, we are able to determine the evolution of the annual seasonality effect over time.

- β_{SA} is the semiannual seasonality parameter,
- $C_{SA}(t_0)$ is the semiannual seasonality centering parameter for t_0. Similar to the definition of $C_A(t_0)$, we define

$$C_{SA}(t_0) = SA_0 + D(t_0), \tag{5.2.4}$$

where the constant time adjustment parameter SA_0 is obtained from a calibration process.

This simple model is considered by several authors [73, 91], although the seasonality feature is handled in a slightly different manner.

Note that the SDE (5.2.1) is different from (4.2.1) in that the seasonality term in (4.2.1) is incorporated into the equilibrium price while the seasonality term in (5.2.1) is proportional to the spot price. Since the natural gas price shows a strong seasonality effect, the seasonality term in (5.2.1) is able to capture the seasonality effect more accurately and thus make a better fit to the market data.

Remark 5.1 (Effect of the Seasonality Term on Gas Price Dynamics)**.** *We can rewrite equation (5.2.1) as*

$$dP = \alpha K_0 dt + (S(t) - \alpha)Pdt + \sigma P dZ. \tag{5.2.5}$$

Since $-(|\beta_A| + |\beta_{SA}|) \leq S(t) \leq |\beta_A| + |\beta_{SA}|$ *according to equation (5.2.2), if*

$$|\beta_A| + |\beta_{SA}| > \alpha, \tag{5.2.6}$$

then there exists certain periods of time within which $S(t) - \alpha > 0$*. In this case, if* P *is large and* $(S(t) - \alpha)Pdt \gg \alpha K_0 dt$ *in equation (5.2.5), then the process (5.2.1) becomes a GBM process with positive drift rate due to the strong seasonality effect. At other times, the process is mean-reverting. Note that the deseasoned process (i.e., setting* $S(t) = 0$

in SDE (5.2.1)) is a mean-reverting process. As indicated in our calibration results in Section 5.3.2, condition (5.2.6) is typically satisfied by the calibrated parameters.

5.2.2 Regime-Switching Model

In order to capture the gas price dynamics more accurately than a one-factor model, [78, 91] propose different two-factor models for the natural gas spot price. In this subsection, we present a one-factor regime-switching model that is able to exhibit behaviour similar to the models introduced in [78, 91] without introducing an additional stochastic factor.

Roughly speaking, our model consists of two regimes; each regime corresponds to a distinct stochastic process (with the same stochastic factor). At any time, the natural gas spot price follows one of these two processes. However, the price process can jump to another regime with some finite probability.

The switch between two regimes can be modeled by a two-state continuous-time Markov chain $m(t)$, taking two values 0 or 1. The value of $m(t)$ indicates the regime in which the risk neutral gas spot price resides at time t. Let $\lambda^{0\to1}dt$ denote the probability of shifting from regime 0 to regime 1 over a small time interval dt, and let $\lambda^{1\to0}dt$ be the probability of switching from regime 1 to regime 0 over dt. Then $m(t)$ can be represented by

$$dm(t) = (1 - m(t-))dq^{0\to1} - m(t-)dq^{1\to0}, \tag{5.2.7}$$

where $t-$ is the time infinitesimally before t, and $q^{0\to1}$ and $q^{1\to0}$ are the independent Poisson processes with intensity $\lambda^{0\to1}$ and $\lambda^{1\to0}$, respectively.

In the regime-switching model, the risk neutral natural gas spot price is modeled by an SDE given by

$$dP = \alpha^{m(t-)}\big(K_0^{m(t-)} - P\big)dt + \sigma^{m(t-)}PdZ + S^{m(t-)}(t)Pdt, \tag{5.2.8}$$

$$S^{m(t-)}(t) = \beta_A^{m(t-)}\sin(2\pi(t - t_0 + C_A(t_0))) + \beta_{SA}^{m(t-)}\sin(4\pi(t - t_0 + C_{SA}(t_0))) \ . \tag{5.2.9}$$

As indicated in equations (5.2.8-5.2.9), within a regime $k \equiv m(t-)$ the gas spot price

follows the process (5.2.1-5.2.2) with parameters $\alpha^k, K_0^k, S^k(t), \sigma^k$ (but the signs of α^k and K_0^k are not constrained). Meanwhile, the stochastic factors for the two regimes are perfectly correlated. Note that we assume that the centering parameters $C_A(t_0)$ and $C_{SA}(t_0)$, as given in equations (5.2.3) and (5.2.4), respectively, are identical for two regimes in order to reduce the number of calibrated parameters.

Remark 5.2 (Mean-Reverting or GBM-Like Process)**.** *From the model (5.2.8-5.2.9), the* deseasoned *spot price in regime $m(t-)$ can follow either a mean-reverting process or a GBM-like process by setting parameter values.*

If we choose $\alpha^{m(t-)} > 0$ and $K_0^{m(t-)} > 0$, then the deseasoned gas price (obtained from setting the seasonality term $S^{m(t-)}(t) = 0$ in SDE (5.2.8)) follows a mean-reverting process

$$dP = \alpha^{m(t-)}(K_0^{m(t-)} - P)dt + \sigma^{m(t-)}PdZ \tag{5.2.10}$$

with equilibrium level $K_0^{m(t-)}$ and mean-reversion rate $\alpha^{m(t-)}$.

If we set $K_0^{m(t-)} = 0$ in equation (5.2.8), then the deseasoned gas price SDE becomes

$$dP = -\alpha^{m(t-)}Pdt + \sigma^{m(t-)}PdZ. \tag{5.2.11}$$

This is a GBM-like process. Specifically, if the drift coefficient $-\alpha^{m(t-)} > 0$, then SDE (5.2.11) is a standard GBM process, i.e., gas price P will drift up at a rate $|\alpha^{m(t-)}|$ at time t; if $-\alpha^{m(t-)} < 0$, then the gas price will drift down at a rate $|\alpha^{m(t-)}|$.

Variations of the Regime-Switching Model

As indicated in Remark 5.2, the deseasoned spot price in each regime can follow either a mean-reverting process or a GBM-like process. Consequently, there exist many possible variations of the regime-switching model by choosing different combinations of the stochastic processes in two regimes. We are interested in the following three variations:

MRMR Variation

The processes in both regimes are mean-reverting with different equilibrium levels, i.e., $K_0^k > 0$, $\alpha^k > 0$, $k \in \{0, 1\}$ in SDE (5.2.8). In this variation, the equilibrium level of the gas spot price switches between two constants, K_0^0, K_0^1, which thus creates a sort of mean-reverting effect on the equilibrium level. This simulates the behaviour of the equilibrium price in the two-factor model proposed by [91], where the gas spot price P follows a one-factor mean-reverting process and its equilibrium price evolves over time according to the other one-factor mean-reverting process.

MRGBM Variation

The process in one regime is mean-reverting while the other regime is a GBM process with a positive drift, i.e., $K_0^0 > 0, K_0^1 = 0, \alpha^0 > 0, \alpha^1 < 0$ in SDE (5.2.8). The mean-reverting regime represents the normal price dynamics, and the GBM regime can be regarded as the sudden drifting up of the gas price driven by exogenous events.

GBMGBM Variation

The processes in both regimes are GBM processes with a positive drift in one regime and a negative drift in the other, i.e., $K_0^0 = K_0^1 = 0, \alpha^0 < 0, \alpha^1 > 0$ in SDE (5.2.8). This simulates the behaviour of the two-factor model in [78], where the risk neutral commodity spot price process is modeled by a GBM-like process given by

$$dP = (r - \delta)Pdt + \sigma PdZ. \tag{5.2.12}$$

Here r is the constant riskless interest rate; δ is the instantaneous convenience yield, following an Ornstein-Uhlenbeck mean-reverting process. The drift coefficient $r - \delta$ can switch between positive and negative values during a time interval since the value of δ is stochastic and may change signs during the interval. Thus the gas price P will either drift up or drift down at any time depending on the sign of $r - \delta$. According to (5.2.11), the GBMGBM variation can produce a behaviour similar to the SDE (5.2.12).

Having presented spot price models, next we calibrate the models to the market gas futures prices and options on futures.

5.3 Calibration to Futures

5.3.1 Data

The data used to test the models consist of monthly observed delivery prices of NYMEX Henry Hub natural gas futures contracts. The data are publicly available on the website `http://www.econstats.com`.

Our data set contains 51 observations in 51 months (one observation each month) during the period from February 2003 to July 2007. Each observation contains delivery prices for the first 14 contracts that correspond to the deliveries in the next 14 consecutive months starting from the month of the observation day.

In order to carry out the calibration, we need to input the gas spot prices. Although there exists a gas spot market in Henry Hub that trades the next day delivery contract, we cannot use the delivery price of the contract as the spot price because the delivery periods for the contracts in the spot market and futures market are different: the delivery lasts for only 24 hours for the former and normally over a whole month for the latter. However, we can regard the delivery price of the next month futures contract, each month on the last trading day of the contract, as the proxy for the gas spot price, since it corresponds to the delivery starting three days later and delivering over the next month[1]. The same approach is used in [57]. Thus, our monthly observation is made on the last trading day of the next month delivery contract[2], where the delivery price for that contract is used as the market spot price during calibration and the delivery prices for the rest of 13 contracts from the observation are used as the market futures prices during calibration

[1] In NYMEX, the trading of the next month delivery contract each month terminates three business days prior to the first calendar day of the next month.

[2] Occasionally, the futures prices on that day are not available on the source website. If that is the case, we use the available price data on the day closest (usually within five days) to the last trading day in the month of the day.

(this amounts to a total of 663 futures prices).

5.3.2 Calibration results

The calibration procedure is given in Appendix E.1. Through calibration to the gas futures contracts, we can obtain all model parameters other than the volatilities. The volatilities will be obtained through calibration to the futures options, as shown in a later section.

As suggested in Section 5.2.2, we are interested in three variations of the regime-switching model. We will determine the best model through calibration, i.e., the model that optimally fits the market data. For this purpose, we set the initial parameter values so that the calibration procedure starts from each of the three variations in Section 5.2.2.

Our calibration results are sensitive to the starting values used in the optimization procedure. For example, if the initial estimates for the parameters has either the MRMR or MRGBM form, the calibrated parameters retain the same form. As we shall see below, good fits to the data can be obtained with either MRMR or MRGBM. However, if we use initial parameters consistent with GBMGBM, then the optimization procedure converges to the MRGBM parameters.

This appears to indicate that the MRMR or MRGBM models are consistent with the market data, while the GBMGBM model does not appear to be consistent with market data. However, we cannot make definite conclusions here, since it is possible that the optimization algorithm may be stuck in a local minimum.

The behaviour of the volatility of the futures price as $T - t$ becomes large depends in general on the calibrated parameters. However, for both regimes in the MRMR model, the volatility of the futures prices declines as the maturity increases. For the MRGBM model, the volatility of the futures prices declines in the MR regime, but the behaviour in the GBM regime is a complicated function of the calibrated parameters and the seasonality terms. It is worthwhile mentioning that the seasonality effects are very large for gas futures prices, and a much larger set of futures prices expiring at long dates would be

needed to determine the long term volatility of the futures prices.

Table 5.1 presents the calibrated risk neutral parameter values for the MR model (5.2.1-5.2.2) and for the MRMR and MRGBM variations. In our calibration procedure, we set a lower bound of $\beta_{SA}^k = 0$. As shown in Table 5.1, the semi-annual seasonality parameters β_{SA}^k for three models are zero, which appears to suggest that a single trigonometric term can satisfyingly approximate the seasonality trend in the futures price data. However we are not sure of this because the result $\beta_{SA}^k = 0$ can be an artifact of the optimization routine (i.e., a local minimum). Meanwhile, the table reveals a strong annual seasonality behaviour under the risk neutral world: condition (5.2.6) is satisfied for the MR model and also for the processes in regime 0 of the MRMR and MRGBM variations. Consequently, Remark 5.1 implies that the corresponding gas price dynamics incorporating the seasonality effect are mean-revering within certain periods of time and switch to (essentially) GBM with positive drift at other times. Regime 1 of the MRMR variation, nevertheless, always shows a mean-reverting effect and that of the MRGBM variation always follows a GBM with positive drift.

From Table 5.1, for the MRMR variation, the equilibrium level in regime 1 is considerably higher than that in regime 0. As a result, regime 0 represents the relatively low price regime and regime 1 represents the relatively high price regime. Similarly, for the MRGBM variation, regime 0 can be regarded as the low price regime and regime 1 represents the regime where the gas price drifts up quickly (according to the value of α^1).

Comparing the calibrated parameter values in regime 0 of the MRMR variation with those in the MR model, we observe that these values are similar except for the equilibrium price: the former has $K_0^0 \approx 4.5$ while the latter has $K_0 \approx 8.7 > K_0^0$. The situation is reversed for regime 1 with $K_0 < K_0^1 \approx 11.7$. The above observation also holds for the MRGBM variation: $K_0^0 < K_0$ in regime 0 and the effective equilibrium price is greater than K_0 in regime 1 (we can imagine that the GBM regime is equivalent to a mean-reverting regime with equilibrium level at $+\infty$).

Note that the risk neutral parameters in Table 5.1 are not necessarily consistent with their counterparts under the real world $\mathbb{P}$ measure. To further illustrate this point, in

Table 5.2 we give, for the MRGBM variation of the regime-switching model, the regimes $\hat{k}(t;\theta)$ where the realized gas spot price resides at various times in our sample, calibrated using the procedure in Section E.1.2. We can observe that the duration of time spent in each regime, implied from Table 5.2, is inconsistent with the risk neutral regime shift intensities $\lambda^{0\to1}$, $\lambda^{1\to0}$ under the $\mathbb{Q}$ measure in Table 5.1: the realized gas price stays in regime 1 for over 45% of the time, while the risk neutral gas price resides at regime 1 for only about 11% of the time. Regime 1 is a regime in which gas price drifts up quickly. The risk neutral price stays in this regime (on average) a much shorter time than the realized price. This observation is consistent with the common paradigm $\mathbb{Q}$ *is more pessimistic than* $\mathbb{P}$, i.e. investors in gas are risk averse and price gas contracts with a pessimistic view of future gas prices.

Table 5.3 provides the dollar and the percentage mean absolute errors between the model and market prices for futures contracts with different delivery months across all observation days. The table illustrates that the MR model performs the worst in terms of both the dollar and the percentage errors. On the other hand, the MRMR and MRGBM variations result in similar errors (with the difference of the overall errors less than 7%), while the MRMR variation outperforms the MRGBM variation for the contracts with relatively long maturities.

Note that these fits were obtained with eleven parameters fitting 663 data points. This fit may not be good enough for trading purposes. However, an *exact fit* can be obtained to any set of futures prices at a given time by adding a time dependent fitting function to the gas price process SDE. However, this fitting function would only have to account for the approximately 7% error obtained from the global calibration, hence would be relatively small. It seems that the overall forward curves for gas can be fit reasonably well with either the MRMR or MRGBM models.

Figure 5.1 illustrates the model implied futures prices and the market prices for the longest maturity contract, which corresponds to the largest calibration errors among all the contracts, in the sample across all observation days starting from February 2003. Figure 5.1a indicate that the MR model fits the market prices poorly in observation days

		MR	MRMR	MRGBM
Parameter	Description	Estimate	Estimate	Estimate
α (α^0)	Mean-reversion rate (for regime 0)	0.406	0.430	0.435
K_0 (K_0^0)	Equilibrium price (for regime 0)	8.678	4.466	4.748
β_A (β_A^0)	Annual seasonality parameter (for regime 0)	0.527	0.600	0.550
β_{SA} (β_{SA}^0)	Semiannual seasonality parameter (for regime 0)	0	0	0
A_0	Annual seasonality time adjustment parameter	0.483	0.441	0.457
α^1	Mean-reversion rate for regime 1		1.033	-0.650
K_0^1	Equilibrium price for regime 1		11.709	0
β_A^1	Annual seasonality parameter for regime 1		0.571	0.555
β_{SA}^1	Semiannual seasonality parameter for regime 1		0	0
$\lambda^{0\to1}$	Intensity of the jump from regime 0 to regime 1		0.304	0.283
$\lambda^{1\to0}$	Intensity of the jump from regime 1 to regime 0		0.975	2.290

TABLE 5.1: *Estimated parameter values for the three models using* 663 *monthly observed futures price data from February 2003 to July 2007. The column MR represents the MR model. The columns MRMR and MRGBM represent the MRMR and MRGBM variation of the regime-switching model, respectively. Units are in terms of years.*

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2003	N/A	0	0	0	0	0	0	0	0	0	0	0
2004	0	0	0	0	0	0	0	1	1	0	0	1
2005	1	1	0	0	N/A	0	0	0	0	0	0	0
2006	1	1	1	1	1	N/A	N/A	1	1	1	1	1
2007	1	1	1	1	1	1	1	N/A	N/A	N/A	N/A	N/A

TABLE 5.2: *Regimes where the realized market gas spot price resides at various times, where the spot price follows the MRGBM variation of the regime-switching model. The Table shows that 29 months correspond to regime 0 and 22 months correspond to regime 1. The N/A in the table corresponds to missing data.*

	Mean absolute error					
Contract maturity	MR	MRMR	MRGBM	MR	MRMR	MRGBM
	In Dollars			In Percentage		
Month+2	0.278	0.240	0.248	3.98	3.57	3.64
Month+3	0.499	0.386	0.388	6.99	5.70	5.56
Month+4	0.645	0.487	0.470	8.96	7.01	6.57
Month+5	0.684	0.504	0.471	9.44	7.08	6.45
Month+6	0.741	0.486	0.502	10.13	6.60	6.60
Month+7	0.827	0.493	0.528	11.14	6.59	6.75
Month+8	0.872	0.492	0.548	11.86	6.66	7.11
Month+9	0.949	0.505	0.563	12.97	6.92	7.37
Month+10	1.011	0.557	0.574	13.93	7.61	7.53
Month+11	1.037	0.603	0.622	14.62	8.20	8.20
Month+12	1.075	0.580	0.640	15.60	8.21	8.53
Month+13	1.118	0.580	0.677	16.65	8.53	9.21
Month+14	1.152	0.585	0.698	17.68	8.97	9.74
Overall	0.838	0.500	0.533	11.84	7.05	7.17

TABLE 5.3: *Mean absolute errors between the model and the market prices for the futures contracts with different delivery months, where the notation Month+k in the first column represents the kth month delivery after the month of the observation day. The errors are given both in dollars and in percentage. The column MR represents the MR model. The columns MRMR and MRGBM represent the MRMR and MRGBM variation of the regime-switching model, respectively.*

close to February 2003. On the contrary, Figures 5.1b and 5.1c show that the MRMR and MRGBM variations of the regime-switching model can reasonably fit the data across all the observation days. Therefore, we conclude that among these models, the regime-switching models outperform the MR model in terms of fitting the market gas forward curves.

5.4 Calibration to Options on Futures

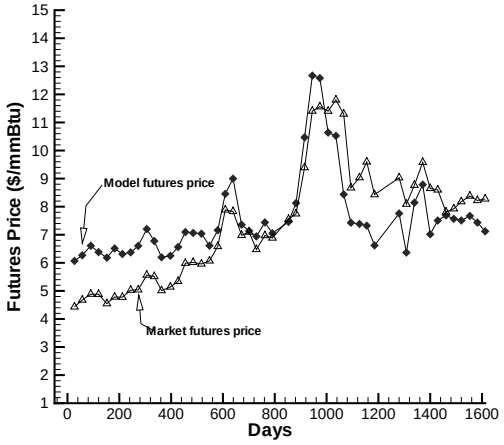
As stated in Remark E.1, the spot price volatilities for the models in Section 5.2 need to be estimated using derivatives other than the futures contracts. Consequently, we calibrate the volatility using market European call options on natural gas futures. The calibration procedure is provided in Appendix E.2.

5.4.1 Calibration Results

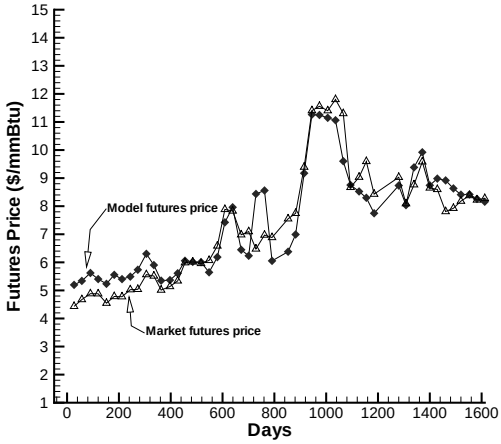
We choose as input the values of twelve European call options from NYMEX on $t = 6/26/2007$ with different strike prices. These options have the same underlying futures contract, which expires in August 2007, denoted by T. The futures price is 7.002 \$/mmBtu at time t. The strike prices with respect to the twelve options range from 6.5 to 7.5 \$/mmBtu, that is, we pick both slightly in the money and slightly out of the money options[3]. We assume that the annual riskless interest rate is $r = 5\%$.

Table 5.4 gives the calibration results and mean absolute errors for the MR model and the MRMR and MRGBM variations of the regime-switching model.

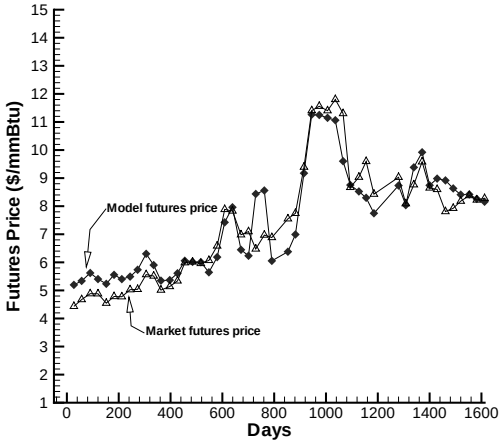
[3]The data set we choose is relatively small. Nevertheless, as an illustration in our simple constant parameter setting, it is sufficient to estimate the volatilities for two regimes. One can add more market data into calibration, such as American options. One can also imagine assuming a volatility surface $\sigma^k = \sigma^k(P, t)$ in model (5.2.8-5.2.9) and calibrating the surface using futures options with different maturities and different strike prices.



(A) MR



(B) MRMR



(C) MRGBM

FIGURE 5.1: *Comparison between the model and the market futures prices for the contract with the longest maturity (for the delivery after 14 months) in the sample across all observation days starting from February 2003. The x-axis represents the number of days between the observation day and the starting date. The model implied prices are computed using the calibrated parameters in Table 5.1. MR represents the MR model. MRMR and MRGBM represent the MRMR and MRGBM variation of the regime-switching model, respectively.*

Model	Volatility		Mean absolute error	
	σ^0	σ^1	In Cents	In Percentage
MR	0.428		1.75	4.27
MRMR	0.406	0.453	0.15	0.59
MRGBM	0.394	0.416	0.15	0.60

TABLE 5.4: *Calibrated volatilities and mean absolute errors for the futures options. The errors are given both in cents and in percentage. The row MR represents the MR model. The rows MRMR and MRGBM represent the MRMR and MRGBM variation of the regime-switching model, respectively.*

5.5 Summary

Our work in this area makes the following contributions:

- We propose a one-factor regime-switching model for the risk neutral natural gas spot price. By adjusting parameter values, the deseasoned process in each regime follows either a mean-reverting process or a geometric Brownian motion (GBM) like process with a positive/negative drift. This produces several variations of the basic model.

- We calibrate model parameters to both market futures and options. We use a two phase calibration process. Under these models, a subset of the parameters can be obtained by calibration to the forward curves. The remaining parameters can be determined by calibration to options. As a result, the computational requirements of the calibration process are reduced compared to more general models. Note that since we are interested in valuation and operation of gas storage, we calibrate to futures and options prices, which gives us the $\mathbb{Q}$ measure parameters directly. This is, of course, distinct from the econometric approach of examining spot price time series, which would generate $\mathbb{P}$ measure parameters.

- Among the three gas price models that we examine, the calibration results show that the MRMR and MRGBM variations of the regime-switching model are capable of fitting the market gas forward curves more accurately than the MR model. The

GBMGBM variation does not appear to be consistent with market data.

Chapter 6

Pricing Natural Gas Storage Contracts under the Regime-Switching Model

In the previous chapter we have showed that the regime-switching model (5.2.8-5.2.9) outperforms the one-factor mean-reverting model (5.2.1-5.2.2) in capturing the dynamics of natural gas futures prices. In this chapter, we apply the calibrated model to price the value of cash flows for a natural gas storage facility. Readers can refer to Chapter 2 for detailed descriptions of the gas storage valuation problem.

This chapter is arranged as follows: we first give the pricing equations for gas storage contracts and discuss the corresponding boundary conditions. We then introduce the numerical scheme for solving the pricing equations and prove the convergence of the scheme to the viscosity solution. Finally, we conduct numerical convergence tests and investigate the optimal operational strategies on storage facilities implied from these gas spot price models.

6.1 Pricing Equation

We use $\hat{V}^k(P,I,t) = \hat{V}(P,I,t,k)$ to represent the value of a natural gas storage facility in regime k when the gas price resides at P, the working gas inventory lies at I and the current time is t. The stochastic control formulation with respect to equation (2.2.8) is

$$\begin{aligned}&\hat{V}^k(P,I,t)\\&= \sup_{c(s)\in C(I(s))} E^{\mathbb{Q}}\left[\int_t^T e^{-r(s-t)}\big[c(s)-a(c(s))\big]P(s)ds + e^{-r(T-t)}\hat{V}^{m(T)}\big(P(T),I(T),T\big)\right],\end{aligned} \tag{6.1.1}$$

where $m(T)$ is the regime where the risk neutral gas spot price resides at time T and other parameters are given in Chapter 2. The operator $E^{\mathbb{Q}}$ is the risk neutral conditional expectation with initial values $P(t) = P$, $I(t) = I$ and $m(t) = k$. Assuming that the risk neutral gas spot price follows the regime-switching model (5.2.8-5.2.9) and following the steps in Appendix A, we can obtain the following coupled HJB equations from the above control equation (6.1.1)

$$\begin{aligned}V^k_\tau = &\frac{1}{2}(\sigma^k)^2P^2V^k_{PP} + \big[\alpha^k(K^k_0 - P) + S^k(t)P\big]V^k_P + \sup_{c\in C(I)}\big\{(c-a(c))P - (c+a(c))V^k_I\big\}\\&- \big(r+\lambda^{k\to(1-k)}\big)V^k + \lambda^{k\to(1-k)}V^{1-k}\ ,\ k\in\{0,1\}\ ,\end{aligned} \tag{6.1.2}$$

where we have changed the variable from $\hat{V}^k(P,I,t)$ to $V^k(P,I,\tau)$ with $\tau = T-t$ and $V^k(P,I,\tau) = \hat{V}^k(P,I,t)$

6.2 Boundary Conditions

In order to completely specify the gas storage problem, we need to provide boundary conditions. As for the terminal boundary conditions, we use the following penalty payoff

function given in (2.2.13):

$$V^k(P, I, \tau = 0) = const. \cdot P \cdot \min\left(I(t = T) - I(t = 0), 0\right) \ , \ k \in \{0, 1\}. \tag{6.2.1}$$

For computational purposes, we truncate the domain from $P \times I \in [0, \infty] \times [0, I_{\max}]$. to a finite computational domain $[0, P_{\max}] \times [0, I_{\max}]$. As $I \to 0$ or $I \to I_{\max}$, according to the arguments in Section 2.2.5, no boundary conditions are needed since the characteristics are outgoing or zero in the I direction.

Taking the limit of equations (6.1.2) as $P \to 0$, we obtain

$$V^k_\tau = \alpha^k K^k_0 V^k_P + \sup_{c \in C(I)} \left\{-(c + a(c))V^k_I\right\} - \left(r + \lambda^{k\to(1-k)}\right)V^k + \lambda^{k\to(1-k)}V^{1-k} \ , \ k \in \{0, 1\} \ . \tag{6.2.2}$$

Since $\alpha^k K^k_0 \geq 0$ for all variations of the regime-switching model we consider (see Section 5.2.2), the characteristics are outgoing in the P direction and we can solve equations (6.2.2) without requiring additional boundary conditions.

As $P \to \infty$, we make the common assumption that $V^k_{PP} \to 0$ (see [87]). We need to deal with one major issue in that the resulting boundary equations require information from outside the computational domain. To see the problem, assuming $V^k_{PP} \to 0$ as $P \to \infty$, then the pricing equations (6.1.2) become

$$\begin{aligned} V^k_\tau = &\left[\alpha^k K^k_0 + \left(S^k(t) - \alpha^k\right)P\right]V^k_P + \sup_{c \in C(I)} \left\{(c - a(c))P - (c + a(c))V^k_I\right\} \\ &- \left(r + \lambda^{k\to(1-k)}\right)V^k + \lambda^{k\to(1-k)}V^{1-k} \ , \ k \in \{0, 1\} \ . \end{aligned} \tag{6.2.3}$$

Using the calibrated parameter values from Table 5.1, we find that $S^0(t) - \alpha^0 > 0$ are positive for certain ranges of t. In this case, the characteristics of equations (6.2.3) are incoming in the P direction at $P \to \infty$ and consequently, a monotone discretization of the equation will require information from outside the computational domain.

This issue can be resolved using the following approximation. The assumption $V^k_{PP} \to$

0 as $P \to \infty$ implies that

$$V^k(P, I, \tau) \approx f^k(I, \tau)P + g^k(I, \tau), \tag{6.2.4}$$

where functions f^k and g^k are independent of P. If we assume that $f^k(I, \tau)P \gg g^k(I, \tau)$ as $P \to \infty$, we can further write

$$V^k \approx f^k(I, \tau)P. \tag{6.2.5}$$

Note that the approximation (6.2.4) is consistent with the payoff (6.2.1). Now the drift term in the boundary equation (6.2.3) can be written as

$$\begin{aligned} \left[\alpha^k K_0^k + \left(S^k(t) - \alpha^k\right)P\right]V_P^k &\approx \left(S^k(t) - \alpha^k\right)PV_P^k \\ &\approx \left(S^k(t) - \alpha^k\right)V^k, \end{aligned} \tag{6.2.6}$$

where the first approximation follows since $\left(S^k(t) - \alpha^k\right)P \gg \alpha^k K_0^k$ as $P \to \infty$ and the second approximation is due to equation (6.2.5). Substituting equation (6.2.6) into equation (6.2.3) results in

$$\begin{aligned} V_\tau^k = &\sup_{c \in C(I)} \left\{(c - a(c))P - (c + a(c))V_I^k\right\} - \left(r + \alpha^k - S^k(t) + \lambda^{k\to(1-k)}\right)V^k \\ &+ \lambda^{k\to(1-k)}V^{1-k}\ ,\ k \in \{0, 1\}. \end{aligned} \tag{6.2.7}$$

Since the drift term in equations (6.2.7) is zero, we are able to provide a monotone discretization for the equation without requiring information from outside the computational domain. (Refer to Section 6.4 for more details.)

6.3 Numerical Scheme

Based on the semi-Lagrangian discretizations in Section 2.3, we can easily derive schemes for solving the gas storage equations (6.1.2) and boundary equations (6.2.1-6.2.2) and (6.2.7) in the regime-switching framework. As demonstrated in Chapter 4, the first-order

fully implicit timestepping scheme is a better choice than the Crank-Nicolson timestepping scheme since the latter does not converge at a higher than first-order rate and cannot guarantee convergence to the viscosity solution of the pricing HJB equation. As a result, in this section we only consider the fully implicit timestepping scheme.

Prior to presenting the scheme, we introduce the notation below that follows a similar manner as that in Chapter 2. We use unequally spaced grids in the P and I directions for the PDE discretization, represented by $[P_0, P_1, \ldots, P_{i_{\max}}]$ and $[I_0, I_1, \ldots, I_{j_{\max}}]$, respectively. We use the discrete timesteps $0 < \Delta\tau <, \ldots, < N\Delta\tau = T$ to discretize the PDEs with $\tau^n = n\Delta\tau$ denoting the nth timestep. We assume that there are mesh size/timestep parameters h satisfying condition (2.3.1).

Let $V^n_{i,j,k}$ denote an approximation of the exact solution $V^k(P_i, I_j, \tau^n)$, where $k \in \{0, 1\}$. Let V^n denote a column vector that includes all values of $V^n_{i,j,k}$ with the index order arranged as $V^n = \left[V^n_{0,0,0}, \ldots, V^n_{i_{\max},0,0}, \ldots, V^n_{0,j_{\max},0}, \ldots, V^n_{i_{\max},j_{\max},1}\right]'$. For future reference, assuming M is a square matrix, then we denote $[MV^n]_{ijk} = (MV^n)_{i,j,k}$, and denote $[MV^n]_{jk}$ as the vector $\left[(MV^n)_{0,j,k}, \ldots (MV^n)_{i_{\max},j,k}\right]'$.

Let $\mathcal{L}$ be a differential operator represented by

$$\mathcal{L}V^k = \frac{1}{2}(\sigma^k)^2 P^2 V^k_{PP} + \left[\alpha^k(K^k_0 - P) + S^k(t)P\right]V^k_P - \left(r + \lambda^{k\to(1-k)}\right)V^k + \lambda^{k\to(1-k)}V^{1-k} \tag{6.3.1}$$

for $P \in [0, P_{\max})$. While $P \to P_{\max}$, according to the boundary equations (6.2.7), we define the operator $\mathcal{L}$ to be

$$\mathcal{L}V^k = -\left(r + \alpha^k - S^k(t) + \lambda^{k\to(1-k)}\right)V^k + \lambda^{k\to(1-k)}V^{1-k}. \tag{6.3.2}$$

The operator $\mathcal{L}$ can be discretized using standard finite difference methods. Let $(\mathcal{L}_h V)^n_{i,j,k}$ denote the discrete value of the operator $\mathcal{L}$ at node (P_i, I_j, τ^n, k). Using central, forward,

or backward differencing in the P direction, we can write

$$
\begin{aligned}
&(\mathcal{L}_h V)^n_{i,j,k} \\
&= \begin{cases} \gamma^n_{i,k} V^n_{i-1,j,k} + \beta^n_{i,k} V^n_{i+1,j,k} - (\gamma^n_{i,k} + \beta^n_{i,k} + r + \lambda^{k\to(1-k)}) V^n_{i,j,k} + \lambda^{k\to(1-k)} V^n_{i,j,1-k} & i < i_{\max}, \\ -\big(r + \alpha^k - S^k(T-\tau^n) + \lambda^{k\to(1-k)}\big) V^n_{i,j,k} + \lambda^{k\to(1-k)} V^n_{i,j,1-k} & i = i_{\max}, \end{cases}
\end{aligned} \tag{6.3.3}
$$

where constants $\gamma^n_{i,k}, \beta^n_{i,k}$ can be determined using an algorithm similar to that given in Appendix B. The algorithm guarantees that $\gamma^n_{i,k}, \beta^n_{i,k}$ satisfy a positive coefficient condition:

$$
\gamma^n_{i,k} \geq 0 \;\; ; \;\; \beta^n_{i,k} \geq 0 \quad i = 0, \ldots, i_{\max} - 1 \; ; \; k = 0, 1 \; ; \; n = 1, \ldots, N. \tag{6.3.4}
$$

Let L^n denote a matrix such that

$$
[L^n V^n]_{ijk} = (\mathcal{L}_h V)^n_{i,j,k}, \tag{6.3.5}
$$

where the discrete value $(\mathcal{L}_h V)^n_{i,j,k}$ is given in equation (6.3.3). We denote by Φ^{n+1} a Lagrange linear interpolation matrix such that

$$
\big[\Phi^{n+1} V^n\big]_{ijk} = V^n_{i,\hat{j},k} + \text{interpolation error}, \tag{6.3.6}
$$

where $V^n_{i,\hat{j},k}$ is an approximation of $V^k(P_i, I^n_{\hat{j}}, \tau^n)$ with $I^n_{\hat{j}}$ given in (2.3.15). Let P denote a column vector satisfying $[P]_i = P_i$. Let ζ^n_{jk} be a diagonal matrix containing control values that need to be determined. Let $a\big(\zeta^n_{jk}\big)$ denote a diagonal matrix with diagonal entries $\big[a\big(\zeta^n_{jk}\big)\big]_{ii} = a\big([\zeta^n_{jk}]_{ii}\big)$. Let I be an identity matrix. Given the above notation, following the matrix form (2.3.25), we can provide the the fully implicit timestepping scheme for the gas storage pricing equations (6.1.2), (6.2.1-6.2.2) and (6.2.7) as follows:

$$
\begin{aligned}
\big[[\mathsf{I} - \Delta\tau L^{n+1}] V^{n+1}\big]_{jk} &= \big[\Phi^{n+1} V^n\big]_{jk} + \Delta\tau\big(\zeta^{n+1}_{jk} - a\big(\zeta^{n+1}_{jk}\big)\big) P, \\
\text{where } \big[\zeta^{n+1}_{jk}\big]_{ii} &= \underset{[\zeta^{n+1}_{jk}]_{ii} \in C^{n+1}_{jk}}{\arg\max} \Big\{ \Big[\big[\Phi^{n+1} V^n\big]_{jk} + \Delta\tau\big(\zeta^{n+1}_{jk} - a\big(\zeta^{n+1}_{jk}\big)\big) P\Big]_i \Big\}
\end{aligned} \tag{6.3.7}
$$

for $j = 0, \ldots, j_{\max}$ and $k = 0, 1$. Here $[\zeta_{jk}^{n+1}]_{ii}$ represents the optimal control in the admissible set C_{jk}^{n+1} (defined in Chapter 2). Again, according to arguments in Chapter 2, we can use max operator, instead of the sup operator, in (6.3.7) since the supremum can be achieved by an admissible control. The discrete optimization problem in (6.3.7) can be solved using the bang-bang and no bang-bang methods as described in Section 2.4.

6.4 Convergence Analysis

In this section we follow the lines in Chapter 3 to prove the convergence of our scheme (6.3.7) to the viscosity solution of pricing equations (6.1.2), (6.2.1-6.2.2) and (6.2.7). Note that the notation of viscosity solutions under the regime-switching setting needs to be slightly modified according to [71]. Briefly, the test functions are defined only for each fixed regime k, and the test functions are above/below the solution only for each fixed k, ignoring the solution value V_j, $j \neq k$. However, as shown in [71], provided a unique and continuous viscosity solution exists, it is sufficient to verify the l_∞-stability, consistency and monotonicity of the scheme in order to guarantee the convergence.

Let us define the matrix

$$Q^{n+1} = \mathsf{I} - \Delta\tau L^{n+1}. \tag{6.4.1}$$

We can write the discrete equations (6.3.7) at each node (P_i, I_j, k) as

$$\begin{aligned}
&\mathcal{G}_{i,j,k}^{n+1}\big(h, V_{i,j,k}^{n+1}, \left\{V_{l,j,k}^{n+1}\right\}_{l\neq i}, V_{i,j,1-k}^{n+1}, \left\{V_{i,j,k}^{n}\right\}\big) \\
&\equiv \min_{[\zeta_{jk}^n]_{ii}\in C_{jk}^{n+1}} \left\{ \left[Q^{n+1}V^{n+1}\right]_{ijk} - \left[\Phi^{n+1}V^{n}\right]_{ijk} - \Delta\tau^n \left[(\zeta_{jk}^{n+1} - a(\zeta_{jk}^{n+1}))P\right]_i \right\} \\
&= 0 \qquad \text{if } n+1 \geq 1,
\end{aligned} \tag{6.4.2}$$

where $\left\{V_{l,j,k}^{n+1}\right\}_{l\neq i}$ is the set of values $V_{l,j,k}^{n+1}$, $l \neq i$, $l = 0, \ldots, i_{\max}$, and $\left\{V_{i,j,k}^{n}\right\}$ is the set of values $V_{i,j,k}^{n}$, $i = 0, \ldots, i_{\max}$, $j = 0, \ldots, j_{\max}$, $k = 0, 1$. We also define $\mathcal{G}_{i,j,k}^{n+1}$ at payoff time

$\tau = 0$ as

$$\mathcal{G}^{n+1}_{i,j,k}\big(h, V^{n+1}_{i,j,k}, \{V^{n+1}_{l,j,k}\}_{l\neq i}, V^{n+1}_{i,j,1-k}, \{V^{n}_{i,j,k}\}\big) \equiv V^{n+1}_{i,j,k} - B(P_i, I_j, k) = 0, \quad \text{if } n+1 = 0, \tag{6.4.3}$$

where $B(P_i, I_j, k)$ is the value of the payoff function (6.2.1) a node (P_i, I_j, k). Consequently, scheme (6.4.2-6.4.3) completely specifies our semi-Lagrangian discretization.

In order to show the convergence, we first prove the following Lemma

Lemma 6.1 (M-matrix property of Q^{n+1})**.** *Assuming that discretization (6.3.3) satisfies the positive coefficient condition (6.3.4) and the timestep condition*

$$\Delta\tau < \frac{1}{S-\alpha} \tag{6.4.4}$$

where

$$S - \alpha \equiv \max\left[\max_{t\in[0,T],k\in\{0,1\}}(S^k(t) - \alpha^k), 0\right] , \tag{6.4.5}$$

then the matrix Q^{n+1} is an M matrix. If $S - \alpha = 0$ in (6.4.5), then constraint (6.4.4) vanishes since $1/(S-\alpha) = \infty$.

Proof. From equations (6.3.3), using conditions (6.3.4) and (6.4.4), we can verify that $-L^{n+1}$ has nonpositive offdiagonal elements and the sum of elements in each row in matrix Q^{n+1} is nonnegative. Note that condition (6.4.4) is needed to show that the above diagonal dominant property holds for the last row of the matrix. Hence Q^{n+1} is an M-matrix. □

Remark 6.2 (Explanation of the Timestep Condition (6.4.4))**.** *Condition (6.4.4) is a mild timestepping constraint since $S^k(t)$ is bounded above by a relatively small constant. For example, using the parameter values from Table 5.1, condition (6.4.4) is equivalent to $\Delta\tau < 5.88$ and $\Delta\tau < 0.83$ for the MRMR and MRGBM variation of the regime-switching model, respectively. This indicates that a timestep of 0.8 year is sufficient to satisfy the condition.*

Based on Lemma 6.1, we obtain the following l_∞-stability result:

Lemma 6.3 (l_∞-stability)**.** *Assuming that discretization (6.4.2-6.4.3) satisfies the conditions required for Lemma 6.1, then as $\Delta\tau \to 0$, scheme (6.4.2-6.4.3) satisfies*

$$\|V^{n+1}\|_\infty \leq D_1\|V^0\|_\infty + D_2, \tag{6.4.6}$$

where D_1, D_2 are bounded constants given by

$$\begin{aligned} D_1 &= \exp\left((S-\alpha)T\right) \\ D_2 &= \begin{cases} \frac{1}{S-\alpha}(D_1-1)\cdot P_{\max}\cdot \max\left\{|C_{\max}(I_{\max})|, |C_{\min}(0)|\right\} & \text{if } S-\alpha > 0, \\ T\cdot P_{\max}\cdot \max\left\{|C_{\max}(I_{\max})|, |C_{\min}(0)|\right\} & \text{if } S-\alpha = 0, \end{cases} \end{aligned} \tag{6.4.7}$$

where $(S-\alpha)$ is defined in (6.4.5). Here $\|V^{n+1}\|_\infty = \max_{i,j,k}|V^{n+1}_{i,j,k}|$. Therefore, according to Definition 3.16, the scheme (6.4.2-6.4.3) is unconditionally l_∞-stable.

Proof. The proof directly follows from using Lemma 6.1 and applying the maximum principle to the discrete equation (6.4.2). We omit the details here. Readers can refer to [40, Theorem 5.5] and [46] for complete stability proofs of the semi-Lagrangian fully implicit scheme for American Asian options and the finite difference schemes for controlled HJB equations, respectively. □

Under the regime-switching setting, the consistency definition in Definition 3.18 needs to be modified according to [71, 6]. Following the proof of Lemma 3.19, we can prove that our scheme (6.4.2-6.4.3) is consistent to the pricing equations (6.1.2), (6.2.1-6.2.2) and (6.2.7) under the regime-switching framework.

Using Lemma 6.1, we can easily verify that the scheme (6.4.2-6.4.3) is monotone, as defined in [11, 6].

From Lemma 6.3 and the discussions above, using the results in [71, 11, 6], we can obtain the following convergence result:

Theorem 6.4 (Convergence to the Viscosity Solution)**.** *Assuming that discretization (6.4.2-6.4.3) satisfies all the conditions required for Lemmas 6.3, and assuming that a unique, continuous viscosity solution exists for equations (6.1.2), (6.2.1-6.2.2) and (6.2.7),*

then scheme (6.4.2-6.4.3) converges to the viscosity solution of gas storage equations (6.1.2), (6.2.1-6.2.2) and (6.2.7).

6.5 Numerical Results

Having introduced the fully implicit semi-Lagrangian discretization scheme in the previous section, this section conducts numerical experiments based on the proposed scheme. Following Chapter 4, we use \$/mmBtu and MMcf as the default units for gas spot price P and gas inventory I, respectively. The numerical experiments use the following nonlinear payoff function as used in Chapter 4:

$$V^k(P, I, \tau = 0) = -2P \max(1000 - I, 0) \ , \ k \in \{0, 1\}. \tag{6.5.1}$$

We first carry out a convergence analysis, assuming that the risk neutral natural gas spot price follows the MRMR variation of the regime-switching model (5.2.8-5.2.9) and taking model parameter values from Tables 5.1 and 5.4. Other input parameters for pricing the value of a gas storage contract are listed in Table 4.1 except that the annual riskless interest rate is set to be $r = 0.05$.

The convergence results for two regimes at the node $(P, I) = (6, 1000)$ obtained from refining the mesh spacing and timestep size are shown in Table 6.1, where we use both the bang-bang and no bang-bang methods for solving the discrete optimization problem in scheme (6.3.7) (or (6.4.2)). The results indicate that the both methods achieve first-order convergence. A similar observation is found for the MRGBM variation.

6.5.1 Optimal Operational Strategies for Different Price Models

Our next step is to investigate the optimal operational strategies implied from the gas spot price models. Figure 6.1a plots the optimal control surface of the MR model (5.2.1-5.2.2) as a function of forward time t and gas price P when $I = 1000$ MMcf. We can verify from the figure that the optimal controls are of the bang-bang type: the controls

P grid nodes	I grid nodes	No. of timesteps	Bang-bang method		No bang-bang method	
			Value	Ratio	Value	Ratio
Regime 0						
87	61	500	3902629	n.a.	4011939	n.a.
173	121	1000	3920017	n.a.	3978633	n.a.
345	241	2000	3932943	1.35	3962706	2.09
689	481	4000	3939135	2.09	3954988	2.06
1377	961	8000	3942166	2.04	3951148	2.01
Regime 1						
87	61	500	4836234	n.a.	4962610	n.a.
173	121	1000	4864240	n.a.	4929533	n.a.
345	241	2000	4881661	1.61	4913620	2.08
689	481	4000	4889176	2.32	4905978	2.08
1377	961	8000	4892715	2.12	4902130	1.99

TABLE 6.1: *The value of a natural gas storage facility in two regimes at $P = 6$ \$/mmBtu and $I = 1000$ MMcf. Risk neutral gas spot price follows MRMR variation of the regime-switching process (5.2.8-5.2.9) with model parameter values given in Tables 5.1 and 5.4. Convergence ratios are presented for the bang-bang and the no bang-bang methods in two regimes. The convergence ratio is defined as the ratio of successive changes in the solution, as the timestep and mesh size are reduced by a factor of two. Constant timesteps are used. The payoff function is given in (6.5.1). Other input parameters are given in Table 4.1 and we use $r = 0.05$. We assume that today is January 1st of the year.*

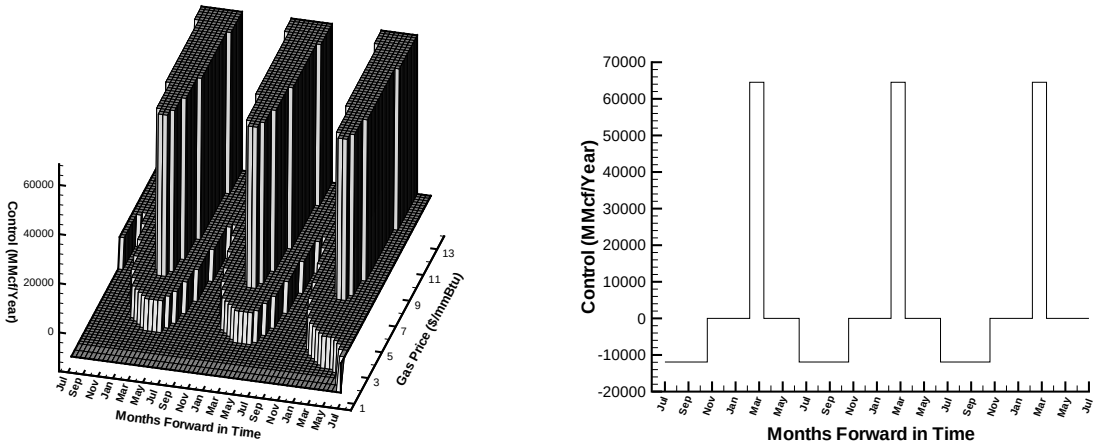
are either producing at the maximum rate $c = c_{\max} > 0$, or injecting at the maximum rate $c = c_{\min} < 0$, or doing nothing with $c = 0$.

Another observation from Figure 6.1a is that for a fixed time t, the control is price dependent; at the same time, the controls evolving over time follow a repeated seasonal pattern. Specifically, in winter months, it is optimal to produce when the price is sufficiently high, to inject when the price is relatively low, and to do nothing when the price lies in between. The higher the price, the longer the production period will be; the lower the price, the longer the injection period will be. Furthermore, the equilibrium level $P = 8.678$ \$/mmBtu approximately resides in the center of the three bang-bang control regions (i.e., the regions of injection, doing nothing and production). This is due to the mean-reverting behaviour of the MR model during the winter period (see the discussions in Remark 5.1 and Section 5.3.2). In summer months, however, the optimal control is to inject or to do nothing, but never to withdraw. The gas price during this period essentially follows a GBM with a positive drift. As such, it is never optimal to withdraw since the price tends to drift up during this period due to the strong seasonality effect. From the discussions above, we conclude that the optimal controls are consistent with the gas price dynamics implied from the calibrated MR model.

We can also observe from Figure 6.1a that the controls converge to zero when $t \to T$ in order to avoid the revenue loss due to the payoff structure (6.5.1).

To see the seasonality effect on the control strategies more clearly, in Figure 6.1b we present the optimal control curve obtained by taking a slice of the control surface in Figure 6.1a at $P = 6$ \$/mmBtu along the t direction. The figure shows that is optimal to produce between February and March (i.e., in the cold season when the gas prices are relatively high), inject between July and October (i.e., in the warm season when the gas prices are relatively low) and do nothing in other seasons.

As a comparison, Figure 6.2 plots, for the MRMR variation of the regime-switching model, the optimal control surface as a function of t and P with $I = 1000$ MMcf and the corresponding control curve at $P = 6$ \$/mmBtu. Comparing three control surfaces in Figures 6.1a, 6.2a and 6.2c, we observe that they have similar seasonal patterns except that



(A) Control surface, MR

(B) Control curve at $P = 6$ \$/mmBtu, MR

FIGURE 6.1: *Optimal control surface for the MR model as a function of forward time t and gas spot price P as well as the corresponding control curve as a function of t when $P = 6$ \$/mmBtu, where the gas inventory resides at $I = 1000$ MMcf. Model parameter values are given in Tables 5.1 and 5.4. Fully implicit timestepping with the no bang-bang method and with constant timesteps is used. Other input parameters are given in Table 4.1 and we use $r = 0.05$. We assume that the starting time $t = 0$ corresponds to July 1st of the year.*

the three bang-bang control regions in each control surface align according to different gas price P, or more precisely, according to the equilibrium prices of the stochastic processes with respect to the three control surfaces. Consequently, the MRMR variation generates different control strategies for two regimes that are consistent with the calibrated processes in those regimes. Moreover, as a regime shift occurs due to an exogenous event, the seasonal control pattern will change accordingly. For example, if the regime shifts from 0 to 1 in March, then the control will switch from producing gas to injecting gas when $P = 6$ \$/mmBtu. As a result, the MRMR variation of the regime-switching model is able to generate controls that reflect the existence of multiple regimes (if we believe this is true) in the market as well as the regime shifts. Therefore, we regard the MRMR variation as a richer model, which has more complex optimal control strategies.

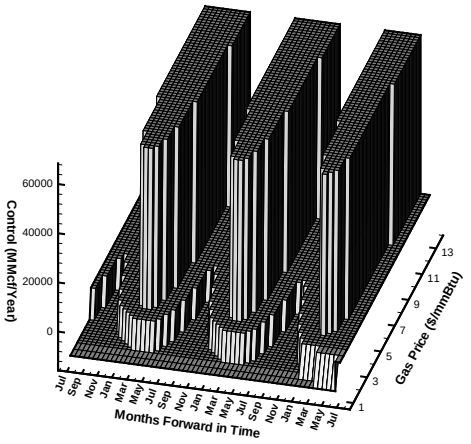
In Figure 6.3, we plot the optimal control surface and the corresponding control curve for the MRGBM variation of the regime-switching model. The control strategies in regime 0 of the MRMR and MRGBM variations are similar. However, in regime 1 of the MRGBM variation, for all gas prices, it is never optimal to produce, even in the winter period (for fixed $I = 1000$). Again, this is in accordance with the GBM behaviour of the gas price process in this regime. Therefore, similar to the MRMR variation, the MRGBM variation also produces regime specific control strategies that are consistent with the gas price dynamics in each regime. Consequently, the MRGBM variation can also produce very different optimal strategies compared to the MR model.

Finally, we note that from a calibration perspective, it is difficult to distinguish between the MRMR or MRGBM models. We would need other evidence to determine whether the gas price dynamics in the high price regime is mean-reverting or GBM.

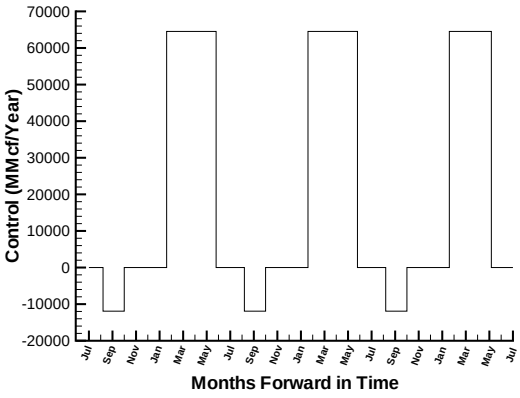
6.6 Summary

Our work in this chapter makes the following contributions:

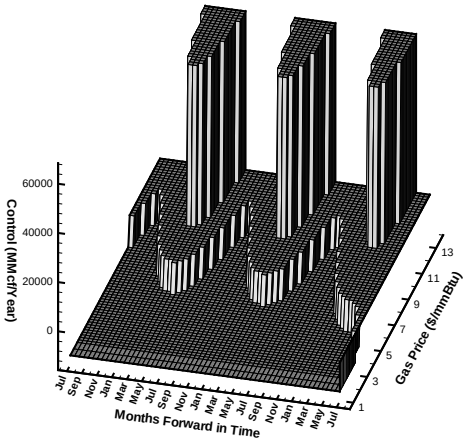
- We extend our fully implicit, semi-Lagrangian timestepping scheme in Chapter 2 for gas storage pricing equation (2.2.12) under a one-factor mean-reverting model



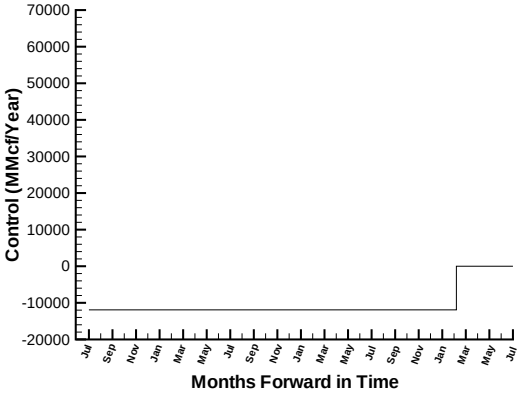
(A) Control surface, Regime 0, MRMR



(B) Control curve at $P = 6$ \$/mmBtu, Regime 0, MRMR

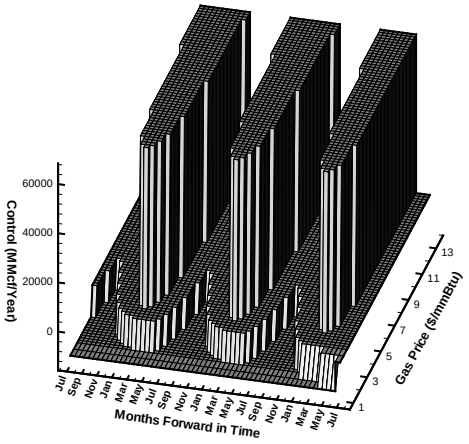


(C) Control surface, Regime 1, MRMR

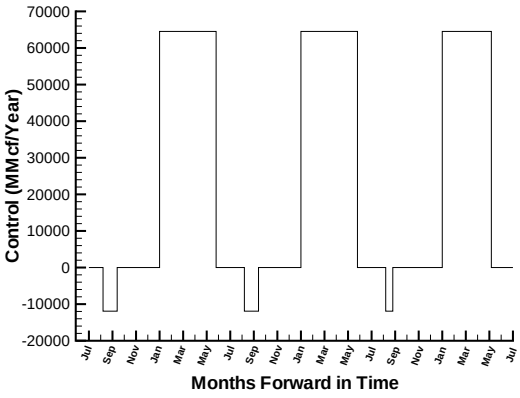


(D) Control curve at $P = 6$ \$/mmBtu, Regime 1, MRMR

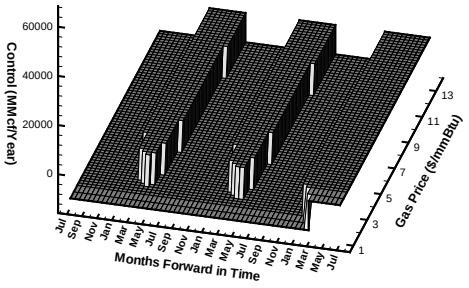
FIGURE 6.2: *Optimal control surface for the MRMR variation of the regime-switching model as a function of forward time t and gas spot price P as well as the corresponding control curve as a function of t when $P = 6$ \$/mmBtu, where the gas inventory resides at $I = 1000$ MMcf. Model parameter values are given in Tables 5.1 and 5.4. Fully implicit timestepping with the no bang-bang method and with constant timesteps is used. Other input parameters are given in Table 4.1 and we use $r = 0.05$. We assume that the starting time $t = 0$ corresponds to July 1st of the year.*



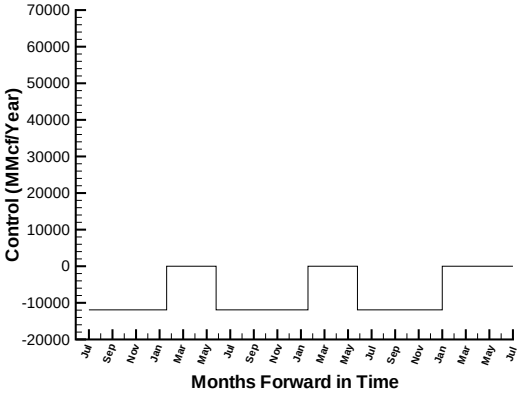
(A) Control surface, Regime 0, MRGBM



(B) Control curve at $P = 6$ \$/mmBtu, Regime 0, MRGBM



(C) Control surface, Regime 1, MRGBM



(D) Control curve at $P = 6$ \$/mmBtu, Regime 1, MRGBM

FIGURE 6.3: *The optimal control surface for the MRGBM variation of the regime-switching model as a function of forward time t and gas spot price P as well as the corresponding control curve as a function of t when $P = 6$ \$/mmBtu, where the gas inventory resides at $I = 1000$ MMcf. Model parameter values are given in Tables 5.1 and 5.4. Fully implicit timestepping with the no bang-bang method and with constant timesteps is used. Other input parameters are given in Table 4.1 and we use $r = 0.05$. We assume that the starting time $t = 0$ corresponds to July 1st of the year.*

to solve the pricing equations (6.1.2) under the regime-switching model, using the model parameters obtained from the calibration. Provided a unique continuous viscosity solution exists, we prove the convergence of the scheme to the viscosity solution of the equations using the results in [71, 11, 6]. The numerical results demonstrate that the scheme converges to the solution at a first-order rate.

- We study the implications of the regime-switching model and the tested one-factor mean-reverting model on the optimal operational strategies for gas storage facilities. Our observations indicate that the regime-switching model, in contrast to one-factor mean-reverting models, is able to produce operational strategies that reflect the existence of multiple regimes in the market as well as the regime shifts due to various exogenous events. Therefore, the regime-switching model is a richer model for pricing the gas storage contracts than the one-factor mean-reverting models.

Chapter 7

Pricing Variable Annuities with a Guaranteed Minimum Withdrawal Benefit (GMWB) under the Discrete Withdrawal Scenario

Variable annuities with GMWBs are extremely popular since these contracts provide investors with the tax-deferred feature of variable annuities as well as the additional benefit of a guaranteed minimum payment. In 2004, sixty-nine percent of all variable annuity contracts sold in the US included a GMWB option [13].

A GMWB contract involves initial payment of a lump sum to an insurance company. This lump sum is then invested in risky assets. The holder of this contract may withdraw funds up to a specified amount each year for the life of the contract, regardless of the performance of the risky assets. As a result, assuming that continuous withdrawals are allowed, the valuation of the GMWB variable annuities is characterized as a stochastic control problem with the withdrawal rate as the control variable.

Prior to pricing continuous withdrawal contracts, in this chapter we formulate a pricing model for more realistic contracts where the withdrawals are allowed only at discrete times [13, 31]. We then present a numerical scheme to solve the pricing model. Based on the

numerical method for discrete withdrawal contracts, in the next chapter we will generalize the approach to value the continuous withdrawal contract.

Using the numerical method in this chapter, in Chapter 9 we will conduct extensive numerical experiments for discrete withdrawal contracts to determine the effect of various parameters and contract specifications on the fair contract value.

7.1 Contract Description

There exist many variations of GMWB variable annuity contracts. In the following, we briefly describe a typical contract that we consider in this thesis. The contract consists of a so called personal sub-account and a virtual guarantee account. The funds in the sub-account are managed by the insurance company investing in a diversified reference portfolio of a specific class of assets. Consequently, the balance in the sub-account is linked to the market performance. At the inception of the policy, the policyholder pays a lump-sum premium to the insurer. This premium forms the initial balance of the sub-account and that of the guarantee account. Prior to the contract maturity, the policyholder is also committed to pay an annual insurance fee proportional to the sub-account balance.

A GMWB option allows the policyholder to withdraw funds from the sub-account at prespecified times (e.g., on an annual or semi-annual basis). Each withdrawal reduces the balance of the guarantee account by the corresponding amount. The policyholder can keep withdrawing as long as the balance of the guarantee account is above zero, even when the sub-account balance falls to zero prior to the policy maturity.

Following [69, 31], we assume the net amount received by the policyholder after a withdrawal is subject to a withdrawal level specified in the contract. If the withdrawal amount does not exceed the contract withdrawal level, then the policyholder receives the complete withdrawal amount. Otherwise, if the withdrawal amount is above the contract level, then the investor receives the remaining amount after a proportional penalty charge is imposed. At the maturity of the policy, the policyholder can choose to receive either the remaining balance of the sub-account if it is positive or the remaining balance of the

guarantee account subject to a penalty charge.

As discussed in [13], for some variations of GMWB contracts, the balance of the guarantee account can increase at certain points in time if no withdrawals have been made so far. In [69, 31] another possibility is discussed whereby an excessive withdrawal may result in a decrease greater than the withdrawal amount in the guarantee account. We will examine some of these complex contract features in a later chapter.

7.2 Discrete Withdrawal Model

In this section we present the pricing model for the valuation of GMWB variable annuities assuming that the investor can withdraw only at discrete times. First, we introduce the notation for the problem. Then we present the pricing equation and the associated boundary conditions.

7.2.1 Problem Notation

We use the following notation for the GMWB variable annuities pricing problem:

- w_0: the premium paid upfront by the policyholder.
- W: the balance of the personal variable annuity sub-account. We have $W = w_0$ at the inception of the contract.
- A: the current balance of the guarantee account. The value of A resides within the interval $[0, w_0]$. We have $A = w_0$ at the inception of the contract.
- $V(W, A, \tau)$: the no-arbitrage value of the variable annuity with GMWB at time $t = T - \tau$ when the value of the sub-account is W and the balance of the guarantee account is A. As usual, we use τ to represent the time to maturity of the contract.
- T: maturity of the policy.
- α: $\alpha \geq 0$, the proportional annual insurance fee paid by the policyholder.

- S: the value of the reference portfolio of assets underlying the variable annuity policy. Following [31], we assume that the risk neutral process of S is modeled by a stochastic differential equation (SDE) given by

$$dS = rSdt + \sigma SdZ, \tag{7.2.1}$$

where $r \geq 0$ is the riskless interest rate, σ is the volatility, dZ is an increment of the standard Gauss-Wiener process.

We denoted by t_O^i, $i = 1, 2, \ldots, K$ the discrete withdrawal times, where $t_O^K = T$, and we denote by $t_O^0 = 0$ the inception time of the policy.

Following [31], we assume there is no withdrawal allowed at $t = 0$ (the inception of the contract). Let $\tau_O^k = T - t_O^i$ be the time to maturity at the ith withdrawal time with $\tau_O^0 = 0$ and $\tau_O^K = T$, where $k = K - i$. In other words, $\tau_O^k, k = 0, \ldots, K-1$ is the kth withdrawal time going backwards in time. Let $\Delta\tau_O^{k+1} = \tau_O^{k+1} - \tau_O^k$.

We denote by γ^k the control variable representing the discrete withdrawal amount at $\tau = \tau_O^k$; γ^k can take any value in $[0, A]$. As such, the dynamics of A are given by

$$\begin{aligned} A(t) &= A(t-) - \gamma^k, \quad \text{if } t = T - \tau_O^k, \\ dA &= 0, \quad \text{otherwise}, \end{aligned} \tag{7.2.2}$$

where $t-$ is the time instantaneously before t.

From (7.2.1) the risk neutral dynamics of W follow an SDE given by

$$dW = (r - \alpha)Wdt + \sigma WdZ + dA, \quad \text{if } W > 0, \tag{7.2.3}$$

$$dW = 0, \quad \text{if } W = 0 \tag{7.2.4}$$

where the dynamics of W are affected by the dynamics of S and A as well as the insurance fee α. Note that the above equations indicate that W will stay at zero from the time it reaches zero.

Let G^k represent the contract withdrawal amount at τ_O^k. If $\gamma^k \leq G^k$, there is no penalty imposed; if $\gamma^k > G^k$, then there is a proportional penalty charge $\kappa(\gamma^k - G^k)$, that is, the net amounted received by the policyholder is $\gamma^k - \kappa(\gamma^k - G^k)$ if $\gamma^k > G^k$, where κ is a positive constant representing the deferred surrender charge. We assume that the penalty (surrender) fees are available to fund the GMWB guarantee.

Consequently, the cash flow received by the policyholder at the discrete withdrawal time $\tau = \tau_O^k$ as a function of γ^k, denoted by $f(\gamma^k)$, is given by

$$f(\gamma^k) = \begin{cases} \gamma^k & \text{if } 0 \leq \gamma^k \leq G^k, \\ \gamma^k - \kappa(\gamma^k - G^k) & \text{if } \gamma^k > G^k. \end{cases} \tag{7.2.5}$$

7.2.2 Pricing Equation

As shown in [31], at the withdrawal time $\tau = \tau_O^k$, V satisfies the following no-arbitrage condition

$$V(W, A, \tau_O^{k+}) = \sup_{\gamma^k \in [0, A]} \left[V\big(\max(W - \gamma^k, 0), A - \gamma^k, \tau_O^k\big) + f(\gamma^k) \right], \quad k = 0, \ldots, K-1, \tag{7.2.6}$$

where τ_O^{k+} denotes the time infinitesimally after τ_O^k.

Within each time interval $[\tau_O^{k+}, \tau_O^{k+1}]$, $k = 0, \ldots, K-1$, the annuity value function $V(W, A, \tau)$, assuming equations (7.2.1-7.2.4), solves the following linear PDE which has A dependence only through equation (7.2.6):

$$V_\tau - \mathcal{L}V = 0, \quad \tau \in [\tau_O^{k+}, \tau_O^{k+1}], \quad k = 0, \ldots, K-1. \tag{7.2.7}$$

where the operator $\mathcal{L}$ is

$$\mathcal{L}V = \frac{1}{2}\sigma^2 W^2 V_{WW} + (r - \alpha) W V_W - rV. \tag{7.2.8}$$

In Appendix I, based on a no-arbitrage hedging argument, we derive the GMWB pricing equation in the presence of a mutual fund management fee. The pricing equation (7.2.6-

7.2.7) can be considered as a special case of the equation in Appendix I with the mutual fund fee set to zero.

7.2.3 Boundary Conditions

We next determine the boundary conditions for equation (7.2.7). Following [31], the terminal boundary condition for the annuity is

$$V(W, A, \tau = 0) = \max(W, (1 - \kappa)A). \tag{7.2.9}$$

This means the policyholder obtains the maximum of the remaining guarantee withdrawal net after the penalty charge ($(1 - \kappa)A$) or the remaining sub-account balance (W).

The domain for equation (7.2.7) is $(W, A) \in [0, \infty] \times [0, w_0]$. For computational purposes, we need to solve the equation in a finite computational domain $[0, W_{\max}] \times [0, w_0]$.

As $A \to 0$, the withdrawal amount γ^k approaches zero. Hence the no-arbitrage condition (7.2.6) reduces to

$$V\big(W, A, \tau_O^{k+}\big) = V\big(W, A, \tau_O^{k}\big), \quad k = 0, \ldots, K - 1, \tag{7.2.10}$$

which means that at the boundary $A = 0$, we only solve the linear PDE (7.2.7) for all $\tau \in [0, T]$.

At $A = w_0$, we simply solve the equations (7.2.6-7.2.7).

At $W = 0$, the no-arbitrage condition (7.2.6) becomes

$$V(0, A, \tau_O^{k+}) = \sup_{\gamma^k \in [0, A]} \left[V\big(0, A - \gamma^k, \tau_O^k\big) + f(\gamma^k) \right], \quad k = 0, \ldots, K - 1. \tag{7.2.11}$$

By taking the limit $W \to 0$, equation (7.2.7) reduces to

$$V_\tau - rV = 0. \tag{7.2.12}$$

We solve equations (7.2.11-7.2.12) at the boundary $W = 0$.

As $W \to \infty$, according to [31], the value function satisfies $V(W, A, \tau) \to e^{-\alpha\tau}W$. As a result, we impose the Dirichlet condition

$$V(W, A, \tau) = e^{-\alpha\tau}W, \quad \text{if } W = W_{\max} \tag{7.2.13}$$

Note that since we will choose $W_{\max} \gg w_0$, evaluating V at $W = W_{\max}$ using equation (7.2.9) gives $V = W_{\max}$, which is the same as evaluating V at $\tau = 0$ using equation (7.2.13).

Let us define solution domains

$$\begin{aligned} \bar{\Omega}_k &= [0, W_{\max}] \times [0, w_0] \times [\tau_O^{k+}, \tau_O^{k+1}] \\ \bar{\Omega} &= \bigcup_k \bar{\Omega}_k = [0, W_{\max}] \times [0, w_0] \times \bigcup_k [\tau_O^{k+}, \tau_O^{k+1}], \quad k = 0, \ldots, K-1. \end{aligned} \tag{7.2.14}$$

The pricing problem for discrete withdrawal contracts can be defined as follows:

Definition 7.1 (Pricing Problem under the Discrete Withdrawal Scenario)**.** The pricing problem for GMWB variable annuities under the discrete withdrawal scenario is defined in $\bar{\Omega}$ as follows: within each domain $\bar{\Omega}_k$, $k = 0, \ldots, K-1$, the solution to the problem is the viscosity solution of a decoupled set of linear PDEs (7.2.7) along the A direction with boundary conditions (7.2.12-7.2.13) and initial condition $V(W, A, \tau_O^{k+})$ computed from the nonlinear algebraic equation (7.2.6).

We next give an auxiliary result and then show that the pricing problem described in Definition 7.1 is well defined in the sense that the solution to the problem is unique.

Lemma 7.2. *If $V(W, A, \tau_O^k)$ is uniformly continuous on $(W, A) \in [0, W_{\max}] \times [0, w_0]$, then $V(W, A, \tau_O^{k+})$ given by equation (7.2.6) is uniformly continuous on $(W, A) \in [0, W_{\max}] \times [0, w_0]$.*

Proof. See Appendix F.1. □

Proposition 7.3. *There exists a unique viscosity solution to the GMWB variable annuity pricing problem described in Definition 7.1. In particular, the solution is continuous on (W, A, τ) within each domain $\bar{\Omega}_k$, $k = 0, \ldots, K-1$.*

Proof. See Appendix F.2 □

Remark 7.4. *We do not define the problem on the continuous region $\tau \in [0, T]$ since the solution can be discontinuous (and hence not well defined) across the observation times τ_O^k, $k = 0, \ldots, K-1$ in the τ direction for fixed (W, A) due to the no-arbitrage condition (7.2.6).*

7.3 Numerical Scheme for the Discrete Withdrawal Model

We use an unequally spaced grid in the W direction for the PDE discretization, represented by $[W_0, W_1, \ldots, W_{i_{\max}}]$ with $W_0 = 0$ and $W_{i_{\max}} = W_{\max}$. Similarly, we use an unequally spaced grid in the A direction denoted by $[A_0, A_1, \ldots, A_{j_{\max}}]$ with $A_0 = 0$ and $A_{j_{\max}} = w_0$. We denote by $0 = \Delta\tau < \ldots < N\Delta\tau = T$ the discrete timesteps. Let $\tau^n = n\Delta\tau$ denote the nth timestep. We assume each discrete withdrawal time τ_O^k coincides with a discrete timestep, denoted by τ^{n_k} with $\tau^{n_0} = \tau^0 = 0$. Let $V(W_i, A_j, \tau^n)$ denote the exact solution of equations (7.2.6-7.2.7) when the value of the variable annuity sub-account is W_i, the guarantee account balance is A_j and discrete time is τ^n. Let $V_{i,j}^n$ denote an approximation of the exact solution $V(W_i, A_j, \tau^n)$.

It will be convenient to define $\Delta W_{\max} = \max_i\big(W_{i+1} - W_i\big)$, $\Delta W_{\min} = \min_i\big(W_{i+1} - W_i\big)$, $\Delta A_{\max} = \max_j\big(A_{j+1} - A_j\big)$, $\Delta A_{\min} = \min_j\big(A_{j+1} - A_j\big)$. We assume that there is a mesh size/timestep parameter h such that

$$\Delta W_{\max} = C_1 h \;\;;\;\; \Delta A_{\max} = C_2 h \;\;;\;\; \Delta\tau = C_3 h \;\;;\;\; \Delta W_{\min} = C_1' h \;\;;\;\; \Delta A_{\min} = C_2' h. \tag{7.3.1}$$

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

As in previous chapters, we use standard finite difference methods to discretize the operator $\mathcal{L}V$ as given in (7.2.8). Let $(\mathcal{L}_h V)_{i,j}^n$ denote the discrete value of the differential operator (7.2.8) at node (W_i, A_j, τ^n). The operator (7.2.8) can be discretized using central,

forward, or backward differencing in the W, A directions to give

$$(\mathcal{L}_h V)^n_{i,j} = \alpha_i V^n_{i-1,j} + \beta_i V^n_{i+1,j} - (\alpha_i + \beta_i + r)V^n_{i,j} \ , \quad i < i_{\max}, \tag{7.3.2}$$

where α_i and β_i are determined using an algorithm in Appendix B. The algorithm guarantees α_i and β_i satisfy the following positive coefficient condition:

$$\alpha_i \geq 0 \quad ; \quad \beta_i \geq 0 \ , \quad i = 0, \ldots, i_{\max} - 1. \tag{7.3.3}$$

At time $\tau = 0$, we apply terminal boundary condition (7.2.9) by

$$V^0_{i,j} = \max(W_i, (1-\kappa)A_j), \quad i = 0, \ldots, i_{\max}, \ j = 0, \ldots, j_{\max}. \tag{7.3.4}$$

At a withdrawal time $\tau^{n_k} = \tau^k_O$, $k = 0, \ldots, K-1$, we apply the no-arbitrage condition (7.2.6) in the following manner. Let $V^n_{\hat{i},\hat{j}}$ be an approximation of $V\big(\max(W_i - \gamma^n_{i,j}, 0), A_j - \gamma^n_{i,j}, \tau^n\big)$ obtained by linear interpolation; in other words, if $\phi(W, A, \tau)$ is a smooth function on (W, A, τ) with $\phi^n_{i,j} = \phi(W_i, A_j, \tau^n)$, then we have

$$\phi^n_{\hat{i},\hat{j}} = \phi\big(\max(W_i - \gamma^n_{i,j}, 0), A_j - \gamma^n_{i,j}, \tau^n\big) + O\Big(\big(\Delta W_{\max} + \Delta A_{\max}\big)^2\Big). \tag{7.3.5}$$

Then at $\tau = \tau^k_O = \tau^{n_k}$, we solve the local optimization problem

$$V^{n+}_{i,j} = \sup_{\gamma^n_{i,j} \in [0, A_j]} \Big[V^n_{\hat{i},\hat{j}} + f\big(\gamma^n_{i,j}\big) \Big], \quad i = 0, \ldots, i_{\max} - 1, \ j = 0, \ldots, j_{\max}, \ n = n_k, \tag{7.3.6}$$

where τ^{n_k+} denotes the time infinitesimally after τ^{n_k}. We describe in Section 7.4 the method used to solve the optimization problem (7.3.6).

Within the interval $\tau \in [\tau^{k+}_O, \tau^{k+1}_O]$, $k = 0, \ldots, K-1$, we use a fully implicit timestep-

ping scheme to discretize (7.2.7). Specifically, we compute $V_{i,j}^{n+1}$ by

$$\begin{aligned}
&V_{i,j}^{n+1} = V_{i,j}^{n+} + \Delta\tau\big(\mathcal{L}_h V\big)_{i,j}^{n+1}, \quad i = 0, \ldots, \ i_{\max} - 1, \ j = 0, \ldots, j_{\max}, \ n+1 = n_k + 1; \\
&V_{i,j}^{n+1} = V_{i,j}^{n} + \Delta\tau\big(\mathcal{L}_h V\big)_{i,j}^{n+1}, \quad i = 0, \ldots, \ i_{\max} - 1, \ j = 0, \ldots, j_{\max}, \ n+1 = n_k + 2, \ldots, n_{k+1}; \\
&V_{i,j}^{n+1} = e^{-\alpha\tau^{n+1}} W_{\max}, \quad i = i_{\max}, \ j = 0, \ldots, j_{\max}, \ n+1 = n_k + 1, \ldots, n_{k+1}.
\end{aligned} \tag{7.3.7}$$

Remark 7.5. *Assuming that* $\max(W_i - \gamma_{i,j}^n, 0)$ *and* $A_j - \gamma_{i,j}^n$ *reside within an interval* $[W_l, W_{l+1}]$ *and* $[A_m, A_{m+1}]$*, respectively, where* $0 \le l < i_{\max}$, $0 \le m < j_{\max}$*, then* $V_{\hat{i},\hat{j}}^n$ *is linearly interpolated using grid nodes* $V_{l,m}^n$, $V_{l+1,m}^n$, $V_{l,m+1}^n$ *and* $V_{l+1,m+1}^n$.

In the discrete equation (7.3.6), $V_{\hat{i},\hat{j}}^n$ *is a function of* $\gamma_{i,j}^n$*, representing the continuous curve on the interpolated surface constructed by linear interpolation using discrete values* $V_{i,j}^n$, $i = 0, \ldots, i_{\max}$, $j = 0, \ldots, j_{\max}$*, along the piecewise line segments* $(W, A)(\gamma_{i,j}^n) = (\max(W_i - \gamma_{i,j}^n, 0), A_j - \gamma_{i,j}^n)$*. Since the values of* $V_{i,j}^n$ *are bounded (see Lemma 7.8), then* $V_{\hat{i},\hat{j}}^n$ *is uniformly continuous on* $\gamma_{i,j}^n$.

According to (7.2.5), $f(\gamma^k)$ *is uniformly continuous on the closed interval* $[0, A_j]$*. Thus the supremum in (7.3.6) is achieved by a control* $\gamma^k \in [0, A_j]$.

7.4 Solution of the Local Optimization Problem

As indicated in (7.3.6), the numerical schemes need to solve a discrete local optimization problem

$$\sup_{\gamma_{i,j}^n \in [0, A_j]} \Big[V_{\hat{i},\hat{j}}^n + f\big(\gamma_{i,j}^n\big) \Big] \tag{7.4.1}$$

at a mesh node (W_i, A_j, τ^n), where $f\big(\gamma_{i,j}^n\big)$ is a piecewise function of $\gamma_{i,j}^n$ given in (7.2.5) and $V_{\hat{i},\hat{j}}^n$ is a function of $\gamma_{i,j}^n$ (see Remark 7.5).

It is expensive to directly solve problem (7.4.1) by constructing the curve $V_{\hat{i},\hat{j}}^n$ and then seeking the maximum of the objective function along the curve. In this section, we present the following consistent approximation to problem (7.4.1). We first select a sequence of control values $\gamma_{i,j}^n$, denoted by $\mathcal{A}_j$, from the interval $[0, A_j]$, where $\mathcal{A}_j$ includes 0, A_j, W_i

and G^k (if $G_r\Delta\tau < A_j$ and $W_i < A_j$), and the distance between two consecutive elements in sequence $\mathcal{A}_j$ is bounded by $O(h)$. We then evaluate the function $V^n_{\hat{i},\hat{j}} + f\big(\gamma^n_{i,j}\big)$ using all elements $\gamma^n_{i,j} \in \mathcal{A}_j$, and return as output the maximum among the set of evaluated values. The above procedure indicates that we actually solve an alternative (and simpler) problem

$$\sup_{\gamma^n_{i,j}\in\mathcal{A}_j} \left[V^n_{\hat{i},\hat{j}} + f\big(\gamma^n_{i,j}\big)\right]. \tag{7.4.2}$$

In terms of a smooth test function, the solutions to problems (7.4.1-7.4.2) satisfy the following conditions:

Proposition 7.6. *Let $\phi(W, A, \tau)$ be a smooth function with $\phi^n_{i,j} = \phi(W_i, A_j, \tau^n)$. Then the optimization procedure introduced above results in*

$$\sup_{\gamma^n_{i,j}\in\mathcal{A}_j} \left[\phi^n_{\hat{i},\hat{j}} + f\big(\gamma^n_{i,j}\big)\right] = \sup_{\gamma^n_{i,j}\in[0,A_j]} \left[\phi^n_{\hat{i},\hat{j}} + f\big(\gamma^n_{i,j}\big)\right] + O(h^2) \tag{7.4.3}$$

$$= \sup_{\gamma^n_{i,j}\in[0,A_j]} \left[\phi\big(\max(W_i - \gamma^n_{i,j}, 0), A_j - \gamma^n_{i,j}, \tau^n\big) + f\big(\gamma^n_{i,j}\big)\right] + O(h^2). \tag{7.4.4}$$

Proof. See Appendix F.3. □

7.5 Convergence of the Numerical Scheme

In this section, we prove the convergence of scheme (7.3.4-7.3.7) to the unique viscosity solution of the pricing problem defined in Definition 7.1 by showing that the scheme is l_∞-stable, pointwise consistent and monotone.

Definition 7.7 (l_∞-Stability)**.** Discretization (7.3.4-7.3.7) is l_∞-stable if

$$\|V^{n+1}\|_\infty \;\leq\; C_4 \;\;, \tag{7.5.1}$$

for $0 \leq n \leq N-1$ as $\Delta\tau \to 0$, $\Delta W_{\min} \to 0$, $\Delta A_{\min} \to 0$, where C_4 is a constant independent of $\Delta\tau$, $\Delta W_{\min}$, $\Delta A_{\min}$. Here $\|V^{n+1}\|_\infty = \max_{i,j} |V^{n+1}_{i,j}|$.

Lemma 7.8 (l_∞ Stability)**.** *If the discretization (7.3.2) satisfies the positive coefficient condition (7.3.3) and linear interpolation is used to compute $V_{\hat{i},\hat{j}}^{n_k}$, then the scheme is stable according to Definition 7.7.*

Proof. The Lemma directly follows from the stability proof of the corresponding scheme under the continuous withdrawal scenario which we discuss in Chapter 8. We refer the reader to the proof of Lemma 8.5 in Chapter 8. □

We can write discrete equations (7.3.7) at a node (W_i, A_j, τ^{n+1}) for $\tau^{n_k+1} \le \tau^{n+1} \le \tau^{n_{k+1}}$ as

$$
\begin{aligned}
&\mathcal{G}_{i,j}^{n+1}\big(h, V_{i,j}^{n+1}, \big\{V_{l,m}^{n+1}\big\}_{\substack{l\neq i\\ m\neq j}}, V_{i,j}^{n+}, \big\{V_{i,j}^{n}\big\}\big)\\
&\equiv \begin{cases}
V_{i,j}^{n+1} - V_{i,j}^{n+} - \Delta\tau\big(\mathcal{L}_h V\big)_{i,j}^{n+1} & \text{if } 0 \le W_i < W_{i_{\max}},\ \ 0 \le A_j \le A_{j_{\max}},\ \ \tau^{n+1} = \tau^{n_k+1};\\
V_{i,j}^{n+1} - V_{i,j}^{n} - \Delta\tau\big(\mathcal{L}_h V\big)_{i,j}^{n+1} & \text{if } 0 \le W_i < W_{i_{\max}},\ \ 0 \le A_j \le A_{j_{\max}},\ \ \tau^{n_k+2} \le \tau^{n+1} \le \tau^{n_{k+1}};\\
V_{i,j}^{n+1} - e^{-\alpha\tau^{n+1}} W_{\max} & \text{if } W_i = W_{i_{\max}},\ \ 0 \le A_j \le A_{j_{\max}},\ \ \tau^{n_k+1} \le \tau^{n+1} \le \tau^{n_{k+1}}
\end{cases}\\
&= 0,
\end{aligned}
\tag{7.5.2}
$$

where $\big\{V_{l,m}^{n+1}\big\}_{\substack{l\neq i\\ m\neq j}}$ is the set of values $V_{l,m}^{n+1}$, $l \neq i$, $l = 0, \ldots, i_{\max}$ and $m \neq j$, $m = 0, \ldots, j_{\max}$, and $\big\{V_{i,j}^{n}\big\}$ is the set of values $V_{i,j}^{n}$, $i = 0, \ldots, i_{\max}$, $j = 0, \ldots, j_{\max}$.

Definition 7.9 (Pointwise Consistency, Discrete Withdrawal)**.** The scheme (7.5.2) is pointwise consistent with the PDE (7.2.7) and boundary conditions (7.2.12-7.2.13) if, for any smooth test function ϕ,

$$
\lim_{h\to 0}\left|\mathcal{G}_{i,j}^{n+1}\big(h, \phi_{i,j}^{n+1}, \big\{\phi_{l,m}^{n+1}\big\}_{\substack{l\neq i\\ m\neq j}}, \phi_{i,j}^{n+}, \big\{\phi_{i,j}^{n}\big\}\big) - (\phi_\tau - \mathcal{L}\phi)_{i,j}^{n}\right| = 0\ , \tag{7.5.3}
$$

for any point in $\bar{\Omega}$, where the solution domain $\bar{\Omega}$ is defined in (7.2.14).

With the above definition, it is straightforward to verify that scheme (7.5.2) is consistent using Taylor series.

Lemma 7.10 (Pointwise Consistency)**.** *The discrete scheme (7.5.2) is pointwise consistent.*

Remark 7.11. *According to Proposition 7.6, the scheme (7.5.2) is still pointwise consistent in the case when the discrete equation (7.3.6) solves the alternative optimization problem (7.4.2).*

The following result shows that scheme (7.5.2) is monotone according to the definition in [11, 6]:

Lemma 7.12 (Monotonicity)**.** *If discretization (7.3.2) satisfies the positive coefficient condition (7.3.3) then discretization (7.5.2) is monotone according to the definition in [11, 6], i.e.,*

$$\begin{aligned}&\mathcal{G}_{i,j}^{n+1}\big(h, V_{i,j}^{n+1}, \big\{X_{l,m}^{n+1}\big\}_{\substack{l\neq i\\ m\neq j}}, X_{i,j}^{n+}, \big\{X_{i,j}^{n}\big\}\big)\\ &\leq \mathcal{G}_{i,j}^{n+1}\big(h, V_{i,j}^{n+1}, \big\{Y_{l,m}^{n+1}\big\}_{\substack{l\neq i\\ m\neq j}}, Y_{i,j}^{n+}, \big\{Y_{i,j}^{n}\big\}\big); \quad \textit{for all } X_{i,j}^{n} \geq Y_{i,j}^{n},\ \forall i,j,n.\end{aligned} \tag{7.5.4}$$

Proof. It is straightforward to verify that the discretization (7.5.2) satisfies inequality (7.5.4) for all mesh nodes (W_i, A_j, τ^n). □

Theorem 7.13 (Convergence to the Viscosity Solution)**.** *Assuming that scheme (7.3.4-7.3.7) satisfies all the conditions required for Lemmas 7.8, 7.10 and 7.12, then as $h \to 0$, scheme (7.3.4-7.3.7) converges to the unique viscosity solution to the pricing problem defined in Definition 7.1 in the domain $\bar{\Omega}$.*

Proof. See Appendix F.4. □

7.6 Summary

Our contribution in this chapter is summarized as follows:

- We formulate a pricing model for the valuation of GMWB variable annuities assuming withdrawals are allowed only at discrete times.
- We present a numerical scheme for solving the pricing model and prove that the scheme converges to the unique viscosity solution of the problem.

- The numerical method proposed in this chapter will be generalized to price GMWB contracts where continuous withdrawals are permitted.

Chapter 8

Pricing GMWB Variable Annuities under the Continuous Withdrawal Scenario

In the previous chapter we have proposed a pricing model for valuing GMWB variable annuities where the investor is allowed to withdraw funds only at discrete times. In this chapter we will study the GMWB valuation problem assuming continuous withdrawals are allowed.

Under the continuous withdrawal scenario, the valuation of the GMWB variable annuities is characterized as a stochastic control problem with the withdrawal rate as the control variable. In contrast to the gas storage valuation problem introduced in Chapter 2, the withdrawal rate at a given time can be either finite or infinite, i.e., the withdrawal rate may be unbounded. A finite withdrawal rate represents a continuous withdrawal, while an infinite rate corresponds to withdrawing a finite *amount* instantaneously.

We model the GMWB variable annuity pricing problem in the continuous withdrawal case as an impulse control problem with two control variables: most of time the contract holder withdraws money continuously at a finite rate (the rate of withdrawal serves as a control variable), and, from time to time, the holder withdraws a finite amount instantaneously (the amount of withdrawal serves as a control variable).

The impulse control formulation has been used in the context of transaction cost models in portfolio optimization [70, 59], liquidity risk and price impact in optimal portfolio selection [83], and execution delay [16]. Refer to [60] for other applications of impulse control in finance and to [72] for a survey of various stochastic controls (including impulse control) and the applications in finance.

Intuitively, the continuous withdrawal problem is the limiting case of the discrete withdrawal problem when the withdrawal intervals decrease to zero. Consequently, we can generalize the numerical scheme for the discrete withdrawal contract in Chapter 7 to solve the impulse control problem corresponding to the continuous withdrawal case. We then prove that the scheme converges to the viscosity solution of the continuous withdrawal problem, provided a strong comparison result holds.

The numerical scheme can be regarded as an extension of the semi-Lagrangian timestepping method for the bounded stochastic control problems (e.g., the gas storage problem described in Chapter 2) to the case of unbounded stochastic control problems. Therefore, we have a unified numerical scheme based on a semi-Lagrangian approach that is able to solve both bounded and unbounded stochastic control problems as well as the discrete cases where the operations are allowed only at discrete times.

At the end of this chapter, we will conduct numerical experiments for the discrete and continuous withdrawal contracts.

8.1 Previous Work

The GMWB variable annuity valuation problem was previously formulated as a singular control problem in [69, 31], where the withdrawal rate is the only control variable. As shown in Remark 8.1, the impulse control formulation, proposed by us, is more general than the singular control formulation.

The authors of [31] use a penalty approach, initially proposed in [47] for pricing American options, to solve the HJB variational inequality for the singular control formulation. Although the penalty method is shown to converge numerically, there is no convergence

proof of the numerical scheme, based on the penalty method, to the viscosity solution of the singular control framework.

The authors of [31] also conduct some experimental computations to show that the numerical solution of the discrete withdrawal contract converges to that of the continuous withdrawal contract as the withdrawal intervals decrease towards zero. Nevertheless, no proof of this convergence was given in [31].

8.2 Continuous Withdrawal Model

Under the continuous withdrawal scenario, we denote by $\hat{\gamma}$ the control variable representing the continuous withdrawal rate. The investor is allowed to withdraw funds from the sub-account at a rate no higher than a contractually specified rate G_r. The investor may withdraw at a rate above G_r, but some penalties are incurred (see below). We denote by $\hat{V}(W, A, t)$ the no-arbitrage value of the GMWB variable annuity at time t when the value of the sub-account is W and the balance of the guarantee account is A.

For continuous withdrawal contracts, the dynamics of the balance of the sub-account W also follows (7.2.3-7.2.4).

In this section, we first recall the singular stochastic control formulation presented in [31], and then propose our impulse control formulation.

8.2.1 Singular Control Formulation

Following [31], we assume $0 \leq \hat{\gamma} \leq \lambda$, where λ is the upper bound of $\hat{\gamma}$. As shown in [31], the dynamics of the balance of the guarantee account A are determined by the dynamics of $\hat{\gamma}$ as follows (in contrast to (7.2.2)):

$$dA = -\hat{\gamma} dt, \tag{8.2.1}$$

where we require that $\hat{\gamma} dt$ be bounded as $dt \to 0$.

Let $\hat{f}(\hat{\gamma})$ be a function of $\hat{\gamma}$ denoting the rate of cash flow received by the policyholder due to the continuous withdrawal. According to [31], we assume that if $\hat{\gamma} \leq G_r$, there is no penalty imposed; if $\hat{\gamma} > G_r$, then there is a proportional penalty charge $\kappa(\hat{\gamma} - G_r)$, that is, the net revenue rate received by the policyholder is $\hat{\gamma} - \kappa(\hat{\gamma} - G_r)$ if $\hat{\gamma} > G_r$, where $\kappa > 0$ is the deferred surrender charge. Consequently, we can write $\hat{f}(\hat{\gamma})$ as a piecewise linear function

$$\hat{f}(\hat{\gamma}) = \begin{cases} \hat{\gamma} & \text{if } 0 \leq \hat{\gamma} \leq G_r, \\ \hat{\gamma} - \kappa(\hat{\gamma} - G_r) & \text{if } \hat{\gamma} > G_r. \end{cases} \tag{8.2.2}$$

Under the singular control framework, the value of $\hat{V}(W, A, t)$ is given by

$$\begin{aligned} &\hat{V}(W, A, t) \\ &= \sup_{\hat{\gamma}(s) \in [0,\lambda]} E^{\mathbb{Q}} \left[\int_t^T e^{-r(s-t)} \hat{f}\big(\hat{\gamma}(s)\big) ds + e^{-r(T-t)} \hat{V}\big(W(T), A(T), T\big) \right], \end{aligned} \tag{8.2.3}$$

where

- $W(s)$ is a path of the balance of the sub-account given by (7.2.3-7.2.4).
- $A(s)$ is a path of the balance of the guarantee account given by (8.2.1).
- $\hat{\gamma}(s)$ is a path of withdrawal rate in the time direction.
- $E^{\mathbb{Q}}$ is the expectation taken under the risk neutral $\mathbb{Q}$ measure conditional on $W(t) = W$ and $A(t) = A$.

Following a procedure similar to Appendix A, the value $\hat{V}(W, A, t)$ satisfies the following HJB equation

$$\hat{V}_t + \mathcal{L}\hat{V} + \sup_{\hat{\gamma} \in [0,\lambda]} \big[\hat{f}\big(\hat{\gamma}\big) - \hat{\gamma}\hat{V}_W - \hat{\gamma}\hat{V}_A \big] = 0, \tag{8.2.4}$$

where the operator $\mathcal{L}$ is given in (7.2.8). Using the notation $V(W, A, \tau)$ with $V(W, A, \tau) = \hat{V}(W, A, T - \tau) = \hat{V}(W, A, t)$, we can rewrite equation (8.2.4) as

$$V_\tau - \mathcal{L}V - \sup_{\hat{\gamma} \in [0,\lambda]} \big[\hat{f}\big(\hat{\gamma}\big) - \hat{\gamma}V_W - \hat{\gamma}V_A \big] = 0. \tag{8.2.5}$$

Since the function $\hat{f}(\hat{\gamma})$ is piecewise linear, the supremum in (8.2.5) is achieved at $\hat{\gamma} = 0$, $\hat{\gamma} = G_r$, or $\hat{\gamma} = \lambda$. Thus, equation (8.2.5) is identical to the following free boundary value problem resulting from evaluating the objective function of the maximization problem at $\hat{\gamma} = 0, G_r, \lambda$, respectively

$$V_\tau - \mathcal{L}V \geq 0, \tag{8.2.6}$$

$$V_\tau - \mathcal{L}V - G_r(1 - V_W - V_A) \geq 0, \tag{8.2.7}$$

$$V_\tau - \mathcal{L}V - \kappa G_r - \lambda\big[(1-\kappa) - V_W - V_A\big] \geq 0, \tag{8.2.8}$$

where the equality holds in at least one of the three cases above. Since $\hat{f}(\hat{\gamma}) = \hat{\gamma}$ for $\hat{\gamma} \in [0, G_r]$, inequalities (8.2.6-8.2.7) are identical to

$$V_\tau - \mathcal{L}V - \sup_{\hat{\gamma}\in[0,G_r]} \big[\hat{\gamma}(1 - V_W - V_A)\big] \geq 0. \tag{8.2.9}$$

Taking the limit $\lambda \to \infty$ (corresponding to an infinite withdrawal rate, or a finite withdrawal amount), inequality (8.2.8) is equivalent to

$$V_W + V_A - (1 - \kappa) \geq 0, \tag{8.2.10}$$

where the expression $V_\tau - \mathcal{L}V - \kappa G_r$ in (8.2.8) becomes negligible as $\lambda \to \infty$.

Consequently, combining inequalities (8.2.9-8.2.10) and using the fact that the equality holds in one of the two cases results in the following HJB variational inequality, as proposed in [31]:

$$\min\left\{V_\tau - \mathcal{L}V - \sup_{\hat{\gamma}\in[0,G_r]} \big(\hat{\gamma} - \hat{\gamma}V_W - \hat{\gamma}V_A\big), V_W + V_A - (1-\kappa)\right\} = 0. \tag{8.2.11}$$

8.2.2 Impulse Control Formulation

As discussed in [92], it is advantageous to reformulate the pricing equation (8.2.11) with a similar HJB variational inequality based on an impulse control argument. Roughly speaking, the policyholder can choose to either withdraw continuously at a rate no greater

than G_r or withdraw a finite amount instantaneously. Withdrawing a finite amount is subject to a penalty charge proportional to the amount of the withdrawal as well as subject to a strictly positive fixed cost, denoted by c. Due to the associated penalty, the withdrawal of a finite amount is optimal only at some discrete stopping times t_s^k.

Since the amount of a finite withdrawal can be infinitesimally small, it is difficult to distinguish two cases: withdrawing at a finite rate or withdrawing an infinitesimal amount. This results in non-uniqueness of the solution to the impulse control formulation. As a result, the nonzero fixed cost c is introduced as a technical tool to distinguish these two cases and resolve the non-uniqueness problem. The nonzero fixed cost is commonly assumed in the impulse control literature [3, 60, 70, 83, 72]. Note that the discrete withdrawal model proposed in Chapter 7 allows the fixed cost to be zero.

Under the impulse control framework, the control of the investor will consist of a combination of a continuous control $\hat{\gamma}(s)$, $\hat{\gamma}(s) \in [0, G_r]$, representing the rate of the continuous withdrawal, and an impulse control (γ^k, t_s^k), $k = 1, 2, \ldots$, representing the amount and time of a withdrawal of a finite amount. Here $t \leq t_s^1 < t_s^2 < \ldots \leq T$ are $\mathcal{F}_s$-stopping times and $\gamma^k \in [0, A(t_s^k-)]$, where $A(t_s^k-)$ is the balance of the guarantee account at the instant infinitesimally before the withdrawal of γ^k occurs.

Given such a control path $\hat{\gamma}(s)$, (γ^k, t_s^k), the dynamics of A satisfies

$$dA = -\hat{\gamma}(s)dt, \quad \text{if } t_s^k \leq s < t_s^{k+1}, \tag{8.2.12}$$

$$A(s) = A(s-) - \gamma^{k+1}, \quad \text{if } s = t_s^{k+1}, \tag{8.2.13}$$

Given (8.2.12-8.2.13), the dynamics of W in (7.2.3-7.2.4) can be rewritten as

$$dW = (r - \alpha)W dt + \sigma W dZ - \hat{\gamma}(s)dt, \quad \text{if } t_s^k \leq s < t_s^{k+1} \text{ and } W > 0, \tag{8.2.14}$$

$$W(s) = \max(W(s-) - \gamma^{k+1}, 0), \quad \text{if } s = t_s^{k+1} \text{ and } W > 0, \tag{8.2.15}$$

$$dW = 0 \quad \text{if } W = 0. \tag{8.2.16}$$

The value of $\hat{V}(W, A, t)$ can be written as the following impulse control problem[1],

[1] To be precise, it is a mixed stochastic control problem. We call it an impulse control problem in

containing both a regular stochastic control (i.e., rate of the continuous withdrawal) and an impulse control (i.e., amount of a finite withdrawal),

$$\hat{V}(W,A,t-) = \sup_{\substack{\hat{\gamma}(s)\in[0,G_r]\\ \gamma^k\in[0,A(t_s^{k-})]\\ t_s^k\in[t,T]}} E^{\mathbb{Q}}\Bigg[\int_t^T e^{-r(s-t)}\hat{\gamma}(s)ds + \sum_{k=1} e^{-r(t_s^k-t)}\Big[(1-\kappa)\gamma^k - c\Big] + e^{-r(T-t)}\hat{V}\big(W(T),A(T),T\big)\Bigg], \tag{8.2.17}$$

where

- $E^{\mathbb{Q}}$ is the expectation taken under the risk neutral $\mathbb{Q}$ measure conditional on $W(t-) = W$ and $A(t-) = A$. The purpose of beginning with $t-$ is to handle the case when $t_s^1 = t$, that is, the first impulse operation occurs at time t.
- The integral term is the total discounted cash due to the continuous withdrawals with the withdrawal rate following the path $\hat{\gamma}(s)$. In this case there is no penalty applied since $\hat{\gamma}(s) \le G_r$.
- The summation term represents the total discounted cash due to the instantaneous withdrawals of finite amounts. Note that a withdrawal of γ^k is always subject to a penalty $\kappa\gamma^k$ and a fixed cost $c > 0$ since it corresponds to the continuous withdrawal with an infinite withdrawal rate.

Using the dynamic programming principle and Itô's Lemma, in Appendix G we heuristically derive the following HJB variational inequality from the impulse control representation (8.2.17):

$$\min\Bigg\{V_\tau - \mathcal{L}V - \sup_{\hat{\gamma}\in[0,G_r]}\big(\hat{\gamma} - \hat{\gamma}V_W - \hat{\gamma}V_A\big), \; V - \sup_{\gamma\in(0,A]}\big[V(\max(W-\gamma,0),A-\gamma,\tau) + (1-\kappa)\gamma - c\big]\Bigg\} = 0, \tag{8.2.18}$$

order to emphasize the impulse operations.

where we change the variable from $\hat{V}(W, A, t)$ to $V(W, A, \tau)$. Here $\hat{\gamma}$ and γ are control variables representing continuous withdrawal rate and finite withdrawal amount, respectively.

In the following, we will consider only the impulse control formulation (8.2.18) with $c > 0$. Although $c > 0$ is required in our theoretical formulation, our numerical scheme proposed in a later section accepts both $c > 0$ and $c = 0$. However, convergence is proved only for the $c > 0$ case. In practice, of course, we would expect that a very small $c > 0$ will have very little effect on the computed solution, and we verify this in our numerical experiments. Indeed, our results for small $c > 0$ are the same (to within discretization errors) as those reported in [31] based on the singular control formulation.

Remark 8.1 (Comparison of Singular Control Formulation and Impulse Control Formulation)**.** *As shown in the following, the impulse control formulation is more general than the singular control formulation:*

- *The singular control formulation requires a piece-wise linear cash flow term in case the control is infinite; the impulse control formulation can be used for any type of cash flow model. For instance, if the revenue function $\hat{f}(\hat{\gamma})$ in (8.2.2) is a nonlinear function of $\hat{\gamma}$ in case $\hat{\gamma} > G_r$, then one cannot obtain an inequality similar to (8.2.10). However, it is straightforward to modify variational inequality (8.2.18) to handle this case.*

- *The singular control formulation allows only zero fixed cost. The impulse control formulation allows both a non-zero and an infinitesimal fixed cost. An infinitesimal fixed cost is effectively the same as the zero fixed cost in the singular control formulation. Indeed, as verified in our numerical experiments, our results for a small fixed cost are the same (to within discretization errors) as the results obtained from the singular control formulation in [31].*

- *It is straightforward to incorporate complex features of real contracts, such as the reset provision on the guarantee level (see [31] and Chapter 9 for details), into*

the impulse control formulation. This will be very difficult for the singular control formulation.

8.2.3 Boundary Conditions for the Impulse Control Problem

In order to completely specify the GMWB variable annuity pricing problem, we need to provide boundary conditions for equation (8.2.18). Similar to the condition (7.2.9) in the discrete withdrawal case, we use the following the terminal boundary condition:

$$V(W, A, \tau = 0) = \max(W, (1 - \kappa)A - c), \tag{8.2.19}$$

where we have incorporated the fixed cost c into the payoff.

The domain for equation (8.2.18) is $(W, A) \in [0, \infty] \times [0, w_0]$. For computational purposes, we need to solve the equation in a finite computational domain $[0, W_{\max}] \times [0, w_0]$.

As $A \to 0$, that is, the guarantee account balance approaches zero, the withdrawal rate $\hat{\gamma}$ must approach zero. Thus by taking $\hat{\gamma} \to 0$ and $A \to 0$ in equation (8.2.18), we obtain a linear PDE

$$V_\tau - \mathcal{L}V = 0 \tag{8.2.20}$$

at $A = 0$. Note that this is essentially a Dirichlet boundary condition at $A = 0$ because we can simply solve equation (8.2.20) independently without using any information other than at $A = 0$.

As $A \to w_0$, since $\hat{\gamma} \geq 0$, the characteristics of the PDE in (8.2.18) are outgoing or zero in the A direction at $A = w_0$. As a result, we can directly solve equation (8.2.18) along the $A = w_0$ boundary, no further information is needed.

As $W \to 0$, following [31], we assume $V_W = 0$ (i.e., W cannot go negative). Taking the limit $W \to 0$ in (8.2.18) and applying $V_W = 0$, we obtain

$$\min \left\{ V_\tau - rV - \sup_{\hat{\gamma} \in [0, G_r]} \left(\hat{\gamma} - \hat{\gamma} V_A \right), V - \sup_{\gamma \in (0, A]} \left[V(0, A - \gamma, \tau) + (1 - \kappa)\gamma - c \right] \right\} = 0 \tag{8.2.21}$$

at $W = 0$. Thus, similar to equation (8.2.20), equation (8.2.21) is essentially a Dirichlet boundary condition since we can solve the equation without requiring any information other than at $W = 0$.

As $W \to \infty$, we apply the following Dirichlet boundary condition as condition (7.2.13):

$$V(W, A, \tau) = e^{-\alpha\tau} W, \quad \text{if } W = W_{\max}. \tag{8.2.22}$$

8.3 Numerical Scheme for the Continuous Withdrawal Model

In this section, we generalize the numerical scheme for the discrete withdrawal contracts in Section 7.3 to solve the HJB variational inequality (8.2.18) and the associated boundary conditions (8.2.19-8.2.22).

The intuition behind this is that the value of a discrete withdrawal contract should converge to that of a continuous withdrawal contract as the observation interval $\Delta\tau_O^k \to 0$. Therefore, we set $\Delta\tau_O^k = \Delta\tau$ in the scheme for the discrete withdrawal contracts, where $k = 0, \ldots, K-1$ and $K = N$. In other words, each discrete timestep τ^n corresponds to a withdrawal time τ_O^k. Then $\Delta\tau_O^k \to 0$ as we take $\Delta\tau \to 0$.

In this case, according to the assumption $\Delta\tau_O^k = \Delta\tau$, the cash flow $f(\gamma_{i,j}^n)$ resulting from (7.2.5) becomes

$$f(\gamma_{i,j}^n) = \begin{cases} \gamma_{i,j}^n & \text{if } 0 \le \gamma_{i,j}^n \le G_r \Delta\tau, \\ \gamma_{i,j}^n - \kappa(\gamma_{i,j}^n - G_r\Delta\tau) - c & \text{if } \gamma_{i,j}^n > G_r\Delta\tau, \end{cases} \tag{8.3.1}$$

where we substitute $G^k = G_r \Delta\tau_O^k = G_r\Delta\tau$ into (7.2.5). We also incorporate the fixed cost c into $f(\gamma_{i,j}^n)$ for any excessive withdrawal above G^k.

We impose condition (8.2.19) at $\tau = 0$ by

$$V_{i,j}^0 = \max(W_i, (1-\kappa)A_j - c)), \quad i = 0, \ldots i_{\max},\ j = 0, \ldots, j_{\max}. \tag{8.3.2}$$

Meanwhile, discrete equations (7.3.6-7.3.7) turn into

$$V_{i,j}^{n+} = \sup_{\gamma_{i,j}^n \in [0,A_j]} \left[V_{\hat{i},\hat{j}}^n + f\left(\gamma_{i,j}^n\right) \right], \quad i = 0, \ldots, i_{\max} - 1, \ j = 0, \ldots, j_{\max}, \tag{8.3.3}$$

$$V_{i,j}^{n+1} = V_{i,j}^{n+} + \Delta\tau \left(\mathcal{L}_h V\right)_{i,j}^{n+1}, \quad i = 0, \ldots, \ i_{\max} - 1, \ j = 0, \ldots, j_{\max}, \tag{8.3.4}$$

$$V_{i,j}^{n+1} = e^{-\alpha\tau^{n+1}} W_{\max}, \quad i = i_{\max}, \ j = 0, \ldots j_{\max} \tag{8.3.5}$$

for $n = 0, \ldots, N-1$, where $\mathcal{L}_h$ is given in (7.3.2). Here $V_{\hat{i},\hat{j}}^n$ is the approximation of $V(\max(W_i - \gamma_{i,j}^n, 0), A_j - \gamma_{i,j}^n, \tau^n)$ by linear interpolation.

Substituting discrete equation (8.3.3) into (8.3.4) gives

$$V_{i,j}^{n+1} - \sup_{\gamma_{i,j}^n \in [0,A_j]} \left[V_{\hat{i},\hat{j}}^n + f\left(\gamma_{i,j}^n\right) \right] - \Delta\tau \left(\mathcal{L}_h V\right)_{i,j}^{n+1} = 0, \quad i = 0, \ldots, i_{\max} - 1, \ j = 0, \ldots, j_{\max}. \tag{8.3.6}$$

Remark 8.2 (Semi-Lagrangian Discretization)**.** *We can also formally obtain scheme (8.3.6) by discretizing the PDE (8.2.5) using the fully implicit, semi-Lagrangian discretization described in Chapter 2, and then taking the limit as $\lambda \to \infty$. Therefore, under the impulse control framework (i.e., when the control is unbounded), the scheme generalized from the discrete withdrawal scenario is identical to that based on the limiting case of a semi-Lagrangian timestepping method. In fact, the correspondence has been verified for the gas storage problem with a bounded control in Chapter 2.*

Remark 8.3. *Since $c > 0$, $f(\gamma_{i,j}^n)$ in (8.3.1) is discontinuous at $G_r\Delta\tau$. Nevertheless, the supremum in (8.3.3) can be achieved by a control $\gamma_{i,j}^n \in [0, A_j]$. To see this, we can write (8.3.3) as*

$$V_{i,j}^{n+} = \max\left\{ \sup_{\gamma_{i,j}^n \in [0, \min(G_r\Delta\tau, A_j)]} \left[V_{\hat{i},\hat{j}}^n + f\left(\gamma_{i,j}^n\right) \right], \sup_{\gamma_{i,j}^n \in (G_r\Delta\tau, A_j]} \left[V_{\hat{i},\hat{j}}^n + f\left(\gamma_{i,j}^n\right) \right] \right\} \tag{8.3.7}$$

with the convention that $(G_r\Delta\tau, A_j] = \emptyset$ if $G_r\Delta\tau \geq A_j$. Since $V_{\hat{i},\hat{j}}^n$ and $f(\gamma_{i,j}^n)$ are continuous on $[0, \min(G_r\Delta\tau, A_j)]$, the first supremum in (8.3.7) can be achieved by a control $\gamma_{i,j}^n \in [0, \min(G_r\Delta\tau, A_j)]$

Equation (8.3.1) implies that

$$f(\gamma_{i,j}^n = G_r\Delta\tau) = \lim_{\gamma_{i,j}^n \to [G_r\Delta\tau]^-} f(\gamma_{i,j}^n) > \lim_{\gamma_{i,j}^n \to [G_r\Delta\tau]^+} f(\gamma_{i,j}^n), \tag{8.3.8}$$

where $\lim_{\gamma_{i,j}^n \to [G_r\Delta\tau]^-} f$ *and* $\lim_{\gamma_{i,j}^n \to [G_r\Delta\tau]^+} f$ *represent the left and right limits of* f *at* $\gamma_{i,j}^n = G_r\Delta\tau$*, respectively. Consequently, if the second supremum in (8.3.7) is achieved by the limiting point* $[G_r\Delta\tau]^+$*, since* $f(G_r\Delta\tau) > f([G_r\Delta\tau]^+)$*, then the value of the first supremum in (8.3.7) will be greater than that of the second one. Thus, the supremum in (8.3.3) can be achieved by a control* $\gamma_{i,j}^n \in [0, A_j]$*.*

Remark 8.4 (Complexity of the Optimization Problem)**.** *The discrete optimization problem in (8.3.6) can be solve using the method described in Section 7.4. Proposition 7.6 still holds in this case.*

Our implementation uses an unequally spaced (W, A) *mesh. As a result, a binary search is required to find the interpolants* $V_{\hat{i},\hat{j}}^n$*. Let us consider the scheme (8.3.2-8.3.5) for the continuous withdrawal case. Since there are* $O(1/h^3)$ *optimizations performed in total (recall that we need to solve a discrete optimization problem (7.4.2) at each mesh node* (W_i, A_j, τ^n) *in this case) and each optimization performs* $O(1/h)$ *linear interpolations (i.e., there are* $O(1/h)$ *elements in sequence* $\mathcal{A}_j$*), resulting in* $O(1/h^4)$ *binary searches (each costing* $O(|log(1/h)|)$*).*

We can reduce the number of binary searches as follows. At each timestep, we transform all the discrete values $V_{i,j}^n$ *in the original unequally spaced* (W, A) *mesh to another equally spaced* (W, A) *mesh by linear interpolation. Then we can solve optimization problems (7.4.2) for all nodes in the equally spaced mesh without using an additional binary search. The above procedure requires only* $O(1/h^3)$ *binary searches in total and results in* $O(h^2)$ *discretization errors for a smooth test function, which hence does not affect the convergence of the numerical scheme to the viscosity solutions. Note that we still require* $O(1/h^4)$ *interpolation operations. An obvious alternative is to use a one dimensional optimization method which would normally not require* $O(1/h)$ *function evaluations at each optimization. However, this is not guaranteed to obtain the global maximum along the curve.*

8.4 Convergence to the Viscosity Solution

As shown in Chapter 3, provided a strong comparison result for the PDE applies, a numerical scheme will converge to the viscosity solution of the equation if it is l_∞-stable, monotone and consistent based on the results in [11, 6]. In this section, we prove the convergence of our numerical scheme (8.3.2-8.3.5) (or scheme (8.3.2), (8.3.5) and (8.3.6)) to the viscosity solution of problem (8.2.18) associated with boundary conditions (8.2.19-8.2.22) by verifying these three properties, assuming a strong comparison principle holds.

8.4.1 l_∞-Stability

At first we show the l_∞-stability of our scheme (8.3.2-8.3.5) by verifying Definition 7.7.

Lemma 8.5 (l_∞-stability)**.** *If the discretization (7.3.2) satisfies the positive coefficient condition (7.3.3) and linear interpolation is used to compute $V^n_{\hat{i},\hat{j}}$, then the scheme (8.3.2-8.3.5) satisfies*

$$\|V^{n+}\|_\infty \le \|V^0\|_\infty + A_{j_{\max}} \quad \textit{and} \quad \|V^n\|_\infty \le \|V^0\|_\infty + A_{j_{\max}} \tag{8.4.1}$$

for $0 \le n \le N$ as $\Delta\tau \to 0$, $\Delta W_{\min} \to 0$, $\Delta A_{\min} \to 0$, where $A_{j_{\max}} = w_0$.

The stability result (8.4.1) also holds for the discrete withdrawal case with $\Delta\tau^n_O > 0$.

Proof. See Appendix H.1. □

8.4.2 Consistency

It will be convenient to rewrite scheme (8.3.2), (8.3.5) and (8.3.6) using the following idea. If $A_j > G_r\Delta\tau$, we can separate the control region into two subregions: $[0, A_j] = [0, G_r\Delta\tau] \cup (G_r\Delta\tau, A_j]$. We will then write equation (8.3.6) in terms of these two subre-

gions. Let us define

$$\mathcal{H}_{i,j}^{n+1}\big(h, V_{i,j}^{n+1}, \big\{V_{l,m}^{n+1}\big\}_{\substack{l\neq i\\ m\neq j}}, \big\{V_{i,j}^{n}\big\}\big) = \frac{1}{\Delta\tau}\Big[V_{i,j}^{n+1} - \sup_{\gamma_{i,j}^n\in[0,\min(A_j, G_r\Delta\tau)]}\big(V_{\hat{i},\hat{j}}^{n} + \gamma_{i,j}^{n}\big) - \Delta\tau\big(\mathcal{L}_h V\big)_{i,j}^{n+1}\Big] \tag{8.4.2}$$

and (assuming $A_j > G_r\Delta\tau$)

$$\begin{aligned}&\mathcal{I}_{i,j}^{n+1}\big(h, V_{i,j}^{n+1}, \big\{V_{l,m}^{n+1}\big\}_{\substack{l\neq i\\ m\neq j}}, \big\{V_{i,j}^{n}\big\}\big)\\ &= V_{i,j}^{n+1} - \sup_{\gamma_{i,j}^n\in(G_r\Delta\tau, A_j]}\Big[V_{\hat{i},\hat{j}}^{n} + (1-\kappa)\gamma_{i,j}^{n} + \kappa G_r\Delta\tau - c\Big] - \Delta\tau\big(\mathcal{L}_h V\big)_{i,j}^{n+1},\end{aligned} \tag{8.4.3}$$

where h is the mesh size/timestep parameter defined in (7.3.1). Note that within (8.4.2-8.4.3), the cash flow term $f(\gamma_{i,j}^n)$ in (8.3.6) is replaced by the piecewise representation given in (8.3.1) based on the subregion where the control $\gamma_{i,j}^n$ resides. Given the definitions of $\mathcal{H}$ and $\mathcal{I}$, we can write scheme (8.3.2), (8.3.5) and (8.3.6) in an equivalent way at a node (W_i, A_j, τ^{n+1}) as

$$\begin{aligned}&\mathcal{G}_{i,j}^{n+1}\big(h, V_{i,j}^{n+1}, \big\{V_{l,m}^{n+1}\big\}_{\substack{l\neq i\\ m\neq j}}, \big\{V_{i,j}^{n}\big\}\big)\\ &\equiv \begin{cases} \mathcal{H}_{i,j}^{n+1} & \text{if } 0\le W_i < W_{i_{\max}},\ 0\le A_j\le G_r\Delta\tau, & 0<\tau^{n+1}\le T;\\ \min\big\{\mathcal{H}_{i,j}^{n+1},\ \mathcal{I}_{i,j}^{n+1}\big\} & \text{if } 0\le W_i < W_{i_{\max}},\ G_r\Delta\tau < A_j\le A_{j_{\max}}, & 0<\tau^{n+1}\le T;\\ V_{i,j}^{n+1} - e^{-\alpha\tau^{n+1}}W_{\max} & \text{if } W_i = W_{i_{\max}},\ 0\le A_j\le A_{j_{\max}}, & 0<\tau^{n+1}\le T;\\ V_{i,j}^{n+1} - & \text{if } 0\le W_i\le W_{i_{\max}},\ 0\le A_j\le A_{j_{\max}}, & \tau^{n+1}=0\\ \max(W_i, (1-\kappa)A_j - c) & \end{cases}\\ &= 0,\end{aligned} \tag{8.4.4}$$

Let $\bar{\Omega} = [0, W_{\max}]\times[0, w_0]\times[0,T]$ be the closed domain in which our problem is defined. The domain $\bar{\Omega}$ can be divided into the following regions:

$$\begin{aligned}&\Omega_{in} = (0, W_{\max})\times(0, w_0]\times(0,T] \ \ ; \ \ \Omega_{W_0} = \{0\}\times(0, w_0]\times(0,T] \ \ ;\\ &\Omega_{A_0} = [0, W_{\max})\times\{0\}\times(0,T] \ \ ; \ \ \Omega_{W_m} = \{W_{\max}\}\times[0, w_0]\times(0,T] \ \ ;\\ &\Omega_{\tau^0} = [0, W_{\max}]\times[0, w_0]\times\{0\},\end{aligned} \tag{8.4.5}$$

where Ω_{in} represents the interior region, and $\Omega_{W_0}, \Omega_{A_0}, \Omega_{W_m}, \Omega_{\tau^0}$ denote the boundary regions. Let us define vector $\mathbf{x} = (W, A, \tau)$, and let $DV(\mathbf{x})$ and $D^2V(\mathbf{x})$ be its first and second derivatives of $V(\mathbf{x})$, respectively. Let us define the following operators:

$$F_{in}\big(D^2V(\mathbf{x}), DV(\mathbf{x}), V(\mathbf{x}), \mathbf{x}\big) = \min\Big\{V_\tau - \mathcal{L}V - \sup_{\hat{\gamma}\in[0,G_r]}\big(\hat{\gamma} - \hat{\gamma}V_W - \hat{\gamma}V_A\big), \\ V - \sup_{\gamma\in(0,A]}\big[V(\max(W-\gamma,0), A-\gamma, \tau) + (1-\kappa)\gamma - c\big]\Big\}, \tag{8.4.6}$$

$$F_{W_0}\big(D^2V(\mathbf{x}), DV(\mathbf{x}), V(\mathbf{x}), \mathbf{x}\big) = \min\Big\{V_\tau - rV - \sup_{\hat{\gamma}\in[0,G_r]}\big(\hat{\gamma} - \hat{\gamma}V_A\big), \\ V - \sup_{\gamma\in(0,A]}\big[V(0, A-\gamma, \tau) + (1-\kappa)\gamma - c\big]\Big\}, \tag{8.4.7}$$

$$F_{A_0}\big(D^2V(\mathbf{x}), DV(\mathbf{x}), V(\mathbf{x}), \mathbf{x}\big) = V_\tau - \mathcal{L}V, \tag{8.4.8}$$

$$F_{W_m}(V(\mathbf{x}), \mathbf{x}) = V - e^{-\alpha\tau}W, \tag{8.4.9}$$

$$F_{\tau^0}(V(\mathbf{x}), \mathbf{x}) = V - \max(W, (1-\kappa)A - c). \tag{8.4.10}$$

Then the pricing problem (8.2.18-8.2.22) can be combined into one equation as follows:

$$F\big(D^2V(\mathbf{x}), DV(\mathbf{x}), V(\mathbf{x}), \mathbf{x}\big) = 0 \quad \text{for all } \mathbf{x} = (W, A, \tau) \in \bar{\Omega}\ , \tag{8.4.11}$$

where F is defined by

$$F = \begin{cases} F_{in}\big(D^2V(\mathbf{x}), DV(\mathbf{x}), V(\mathbf{x}), \mathbf{x}\big) & \text{if } \mathbf{x} \in \Omega_{in}, \\ F_{W_0}\big(D^2V(\mathbf{x}), DV(\mathbf{x}), V(\mathbf{x}), \mathbf{x}\big) & \text{if } \mathbf{x} \in \Omega_{W_0}, \\ F_{A_0}\big(D^2V(\mathbf{x}), DV(\mathbf{x}), V(\mathbf{x}), \mathbf{x}\big) & \text{if } \mathbf{x} \in \Omega_{A_0}, \\ F_{W_m}\big(V(\mathbf{x}), \mathbf{x}\big) & \text{if } \mathbf{x} \in \Omega_{W_m}, \\ F_{\tau^0}\big(V(\mathbf{x}), \mathbf{x}\big) & \text{if } \mathbf{x} \in \Omega_{\tau_0}. \end{cases} \tag{8.4.12}$$

In order to demonstrate consistency as defined in [11, 6], we first need some intermediate results. We define operators

$$F_{A'}\big(D^2V(\mathbf{x}), DV(\mathbf{x}), V(\mathbf{x}), \mathbf{x}\big) = V_\tau - \mathcal{L}V - \sup_{\hat{\gamma}\in[0,A/\Delta\tau]}\big(\hat{\gamma} - \hat{\gamma}V_W - \hat{\gamma}V_A\big),$$

$$\text{where } 0 \le A/\Delta\tau \le G_r \ , \tag{8.4.13}$$

$$F_{W'}\big(D^2V(\mathbf{x}), DV(\mathbf{x}), V(\mathbf{x}), \mathbf{x}\big) = V_\tau - rV - \sup_{\hat{\gamma}\in[0,A/\Delta\tau]} \big(\hat{\gamma} - \hat{\gamma}V_A\big), \quad \text{where } 0 \le A/\Delta\tau \le G_r. \tag{8.4.14}$$

Lemma 8.6. *Let* $\mathbf{x} = (W_i, A_j, \tau^{n+1})$. *Suppose the mesh size and the timestep parameter satisfy conditions (7.3.1) and assume*

$$\Delta W_{\min} \ge G_r \Delta\tau. \tag{8.4.15}$$

Then for any smooth function $\phi(W, A, \tau)$ *having bounded derivatives of all orders in* $(W, A, \tau) \in \bar{\Omega}$, *with* $\phi_{i,j}^{n+1} = \phi(W_i, A_j, \tau^{n+1})$, *and for* h *sufficiently small, we have that*

$$\mathcal{G}_{i,j}^{n+1}\big(h, \phi_{i,j}^{n+1} + \xi, \big\{\phi_{l,m}^{n+1} + \xi\big\}_{\substack{l\neq i\\ m\neq j}}, \big\{\phi_{i,j}^{n} + \xi\big\}\big)$$

$$= \begin{cases} F_{in} + O(h) + c(\mathbf{x})\xi & \text{if } 0 < W_i < W_{i_{\max}}, \ G_r\Delta\tau < A_j \le A_{j_{\max}}, \ 0 < \tau^{n+1} \le T; \\ F_{W_0} + O(h) + c(\mathbf{x})\xi & \text{if } W_i = 0, \ G_r\Delta\tau < A_j \le A_{j_{\max}}, \ 0 < \tau^{n+1} \le T; \\ F_{W'} + O(h) + c(\mathbf{x})\xi & \text{if } W_i = 0, \ 0 < A_j \le G_r\Delta\tau, \ 0 < \tau^{n+1} \le T; \\ F_{A_0} + O(h) + c(\mathbf{x})\xi & \text{if } 0 \le W_i < W_{i_{\max}}, \ A_j = 0, \ 0 < \tau^{n+1} \le T; \\ F_{A'} + O(h) + c(\mathbf{x})\xi & \text{if } 0 < W_i < W_{i_{\max}}, \ 0 < A_j \le G_r\Delta\tau, \ 0 < \tau^{n+1} \le T; \\ F_{W_m} + c(\mathbf{x})\xi & \text{if } W_i = W_{i_{\max}}, \ 0 \le A_j \le A_{j_{\max}}, \ 0 < \tau^{n+1} \le T; \\ F_{\tau^0} + c(\mathbf{x})\xi & \text{if } 0 \le W_i \le W_{i_{\max}}, \ 0 \le A_j \le A_{j_{\max}}, \ \tau^{n+1} = 0, \end{cases} \tag{8.4.16}$$

where ξ *is a constant,* $c(\mathbf{x})$ *is a bounded function of* $\mathbf{x}$ *satisfying* $|c(\mathbf{x})| \le \max(r, 1)$ *for all* $\mathbf{x} \in \bar{\Omega}$, *operators* $F_{in}, F_{W_0}, F_{A_0}, F_{A'}, F_{W'}$ *are functions of* $(D^2\phi(\mathbf{x}), D\phi(\mathbf{x}), \phi(\mathbf{x}), \mathbf{x})$, *and operators* F_{W_m}, F_{τ^0} *are functions of* $(\phi(\mathbf{x}), \mathbf{x})$.

Proof. See Appendix H.2. □

Remark 8.7. *To ease the presentation of the scheme, we impose the grid size condition*

(8.4.15) for the purpose of making

$$V\big(\max(W_i-\gamma^n_{i,j},0),A_j-\gamma^n_{i,j},\tau^n\big)=V\big(W_i-\gamma^n_{i,j},A_j-\gamma^n_{i,j},\tau^n\big),\quad \forall\gamma^n_{i,j}\in[0,G_r\Delta\tau]\ \textit{and}\ \forall W_i>0. \tag{8.4.17}$$

However, we can avoid this condition by modifying the scheme according to the following ideas: at first we extend the W grid in the $W<0$ direction, that is, the extended grid includes nodes with negative W values. Then at each timestep $\tau^{n+1}>0$, we first compute $V^{n+1}_{0,j}$ at $W=0$ using discrete equation (8.3.6) (this is possible since we do not require information from other grid nodes in W direction), and then we set $V^{n+1}_{i,j}=V^{n+1}_{0,j}$ for all $W_i<0$. Finally, we compute $V^{n+1}_{i,j}$ using a modification of equation (8.3.6):

$$V^{n+1}_{i,j}-\sup_{\gamma^n_{i,j}\in[0,A_j]}\Big[V^n_{\hat{i},\hat{j}}+f\big(\gamma^n_{i,j}\big)\Big]-\Delta\tau\big(\mathcal{L}_hV\big)^{n+1}_{i,j}=0, \tag{8.4.18}$$

where the term $V^n_{\hat{i},\hat{j}}$ is the approximation of $V(W_i-\gamma^n_{i,j},A_j-\gamma^n_{i,j},\tau^n)$ by linear interpolation. Since $V^n_{\hat{i},\hat{j}}$ exists in the case when $W_i-\gamma^n_{i,j}<0$ and is equal to the approximation of $V(0,A_j-\gamma^n_{i,j},\tau^n)$, the modified scheme is identical to the original one. Therefore, with respect to the modified scheme, (8.4.17) follows without imposing condition (8.4.15). Hence condition (8.4.15) can be eliminated.

Remark 8.8. *It can be verified that the operators $F_{in}(M,p,g,\mathbf{x})$, $F_{W_0}(M,p,g,\mathbf{x})$, $F_{A_0}(M,p,g,\mathbf{x})$, $F_{A'}(M,p,g,\mathbf{x})$, and $F_{W'}(M,p,g,\mathbf{x})$ defined in (8.4.6-8.4.8) and (8.4.13-8.4.14) are continuous on $(M,p,g,\mathbf{x})$, given a smooth function $g(\mathbf{x})$; meanwhile, operators $F_{W_m}(g,\mathbf{x})$ and $F_{\tau^0}(g,\mathbf{x})$ in (8.4.9-8.4.10) are continuous on $(g,\mathbf{x})$. In particular, $\phi-\sup_{\gamma\in(0,A]}\big[\phi(\max(W-\gamma,0),A-\gamma,\tau)+(1-\kappa)\gamma-c\big]$ is continuous on $\mathbf{x}$ based on an argument similar to the proof of Lemma 7.2.*

The Lemma below proves that scheme (8.4.4) is consistent, as defined in [11, 6].

Lemma 8.9 (Consistency)**.** *Assuming all the conditions in Lemma 8.6 are satisfied, then the scheme (8.4.4) is consistent to the impulse control problem (8.2.18-8.2.22) in $\bar{\Omega}$ according to the definition in [11, 6]. That is, for all $\hat{\mathbf{x}}=(\hat{W},\hat{A},\hat{\tau})\in\bar{\Omega}$ and any function $\phi(W,A,\tau)$ having bounded derivatives of all orders in $(W,A,\tau)\in\bar{\Omega}$ with $\phi^{n+1}_{i,j}=$*

$\phi(W_i, A_j, \tau^{n+1})$ *and* $\mathbf{x} = (W_i, A_j, \tau^{n+1})$*, we have*

$$\limsup_{\substack{h\to 0\\ \mathbf{x}\to\hat{\mathbf{x}}\\ \xi\to 0}} \mathcal{G}_{i,j}^{n+1}\big(h, \phi_{i,j}^{n+1}+\xi, \{\phi_{l,m}^{n+1}+\xi\}_{\substack{l\neq i\\ m\neq j}}, \{\phi_{i,j}^{n}+\xi\}\big) \le F^*\big(D^2\phi(\hat{\mathbf{x}}), D\phi(\hat{\mathbf{x}}), \phi(\hat{\mathbf{x}}), \hat{\mathbf{x}}\big), \tag{8.4.19}$$

and

$$\liminf_{\substack{h\to 0\\ \mathbf{x}\to\hat{\mathbf{x}}\\ \xi\to 0}} \mathcal{G}_{i,j}^{n+1}\big(h, \phi_{i,j}^{n+1}+\xi, \{\phi_{l,m}^{n+1}+\xi\}_{\substack{l\neq i\\ m\neq j}}, \{\phi_{i,j}^{n}+\xi\}\big) \ge F_*\big(D^2\phi(\hat{\mathbf{x}}), D\phi(\hat{\mathbf{x}}), \phi(\hat{\mathbf{x}}), \hat{\mathbf{x}}\big), \tag{8.4.20}$$

where F^* *and* F_* *are respectively the usc and lsc envelopes of* F*, as defined in Definition 3.7.*

Proof. See Appendix H.3. □

8.5 Monotonicity

It is straightforward to verify that scheme (8.4.4) is monotone. We omit the proof.

Lemma 8.10 (Monotonicity)**.** *If the discretization (7.3.2) satisfies the positive coefficient condition (7.3.3) and linear interpolation is used to compute* $V_{\hat{i},\hat{j}}^n$*, then the discretization (8.4.4) is monotone according to the definition*

$$\begin{aligned}&\mathcal{G}_{i,j}^{n+1}\big(h, V_{i,j}^{n+1}, \{X_{l,m}^{n+1}\}_{\substack{l\neq i\\ m\neq j}}, \{X_{i,j}^{n}\}\big)\\ &\le \mathcal{G}_{i,j}^{n+1}\big(h, V_{i,j}^{n+1}, \{Y_{l,m}^{n+1}\}_{\substack{l\neq i\\ m\neq j}}, \{Y_{i,j}^{n}\}\big); \quad \textit{for all } X_{i,j}^n \ge Y_{i,j}^n,\ \forall i,j,n.\end{aligned} \tag{8.5.1}$$

8.6 Convergence

In order to prove the convergence of our scheme using the results in [11, 6], similar to Assumption 3.15, we need to assume the following strong comparison result, as defined in [11, 6], for equation (8.2.18).

Assumption 8.11. If u and v are an usc subsolution and a lsc supersolution of the pricing equation (8.2.18) associated with the boundary conditions (8.2.19-8.2.22), respectively, then

$$u \leq v \quad \text{on } \Omega_{in}. \tag{8.6.1}$$

The strong comparison result is proved for other similar (but not identical) impulse control problems in [3, 83, 70, 56]. From Lemmas 8.5, 8.9, 8.10 and Assumption 8.11, using the results in [11, 6], we can obtain the following convergence result:

Theorem 8.12 (Convergence to the Viscosity Solution)**.** *Assuming that discretization (8.3.2-8.3.5) (or scheme (8.3.2), (8.3.5), (8.3.6), or scheme (8.4.4)) satisfies all the conditions required for Lemmas 8.5, 8.9 and 8.10, and that Assumption 8.11 is satisfied, then scheme (8.3.2-8.3.5) converges to the unique continuous viscosity solution of the problem (8.2.18-8.2.22) in* Ω_{in}.

Remark 8.13 (Domain of Convergence)**.** *Note that we only consider convergence in* Ω_{in}. *As discussed in [83], in general, the strong comparison result may only hold in* Ω_{in} *for impulse control problems.*

8.7 Numerical Experiments

Having presented a consistent numerical scheme for pricing the GMWB variable annuities in Chapter 7 and this chapter, respectively assuming the discrete and continuous withdrawal scenarios, in this section we conduct numerical experiments based on the scheme.

Under the continuous withdrawal scenario, we observe that the numerical solutions obtained by choosing a sufficiently small fixed cost (e.g., $c = 10^{-8}$) are identical to those obtained by choosing $c = 0$ up to at least seven digits. Since the solutions are also close to that given in [31] (see, e.g., Table 8.4), this suggests that our impulse control formulation (8.2.18) will converge to the singular control formulation (8.2.11) as $c \to 0$. It also shows that our scheme can solve both the singular control problem (8.2.11) with $c = 0$ and

Parameter	Value
Expiry time T	10.0 years
Interest rate r	.05
Maximum withdrawal rate G_r	10/year
Withdrawal penalty κ	.10
Initial Lump-sum premium w_0	100
Initial guarantee account balance	100
Initial sub-account value	100

TABLE 8.1: *Common data used in the numerical tests.*

the impulse control problem (8.2.18) with $c > 0$. We will use $c = 10^{-8}$ in the numerical experiments below.

Recall that the computational domain has been localized in the W direction to $[0, W_{\max}]$. Initially, we set $W_{\max} = 1000$. We repeated the computations with $W_{\max} = 5000$. All the numerical results at $t = 0, A = W = w_0$ were the same to seven digits. In the following, all the results are reported with $W_{\max} = 1000$.

Table 8.1 gives the common input parameters for the numerical tests in this section. We first carry out a convergence analysis for the GMWB guarantees with the mesh size/timestep parameters chosen in Table 8.2. Table 8.3 presents the convergence results for the value of the GMWB guarantee with respect to two volatility values, assuming a zero insurance fee and continuous withdrawal. The convergence ratio in the table is defined as the ratio of successive changes in the solution, as the timestep and mesh size are reduced by a factor of two. A ratio of two indicates first-order convergence. As shown in Table 8.3, our scheme achieves a first-order convergence as the convergence ratios are approximately two. The table also reveals that a greater volatility produces a higher contract value.

Since no fee is paid at the inception of a GMWB contract, the insurance company needs to charge a proportional insurance fee α so that the contract value V is equal to the initial premium w_0 paid by the investor. This is the no-arbitrage or *fair* fee. That is, let $V(\alpha; W = w_0, A = w_0, t = 0)$ be the value of a GMWB contract at the contract inception as a function of α. Then the fair insurance fee is a solution to the algebraic equation

Level	W Nodes	A Nodes	Timesteps
0	65	51	60
1	129	101	120
2	257	201	240
3	513	401	480
4	1025	801	960
5	2049	1601	1920

TABLE 8.2: *Grid and timestep data for convergence tests.*

Refinement level	$\sigma = .20$		$\sigma = .30$	
	Value	Ratio	Value	Ratio
1	107.6950	n.a.	115.8032	n.a.
2	107.7132	n.a.	115.8457	n.a.
3	107.7232	1.82	115.8678	1.92
4	107.7284	1.92	115.8787	2.03
5	107.7313	1.79	115.8842	1.98

TABLE 8.3: *Convergence study for the value of the GMWB guarantee at $t = 0$, $W = A = w_0 = 100$. No insurance fee ($\alpha = 0$) is imposed. Data are given in Table 8.1. Continuous withdrawal is permitted.*

$V(\alpha; w_0, w_0, 0) = w_0$. In this thesis, we solve the equation numerically using Newton iteration. Our experiment indicates that Newton iteration seems able to converge to a unique solution at a close-to-second-order rate.

Table 8.4 shows the convergence of the fair insurance fees assuming continuous withdrawal for two volatility values $\sigma = .2$ and $\sigma = .3$. Table 8.4 also lists the corresponding fees computed in [31]. These results are close to those reported in [31].

Table 8.5 computes the fair insurance fees under the discrete withdrawal scenario with withdrawal interval being half a year and one year, respectively. Comparing Tables 8.4 and 8.5, we find that the insurance fees increases as the specified withdrawal frequency increases (from once every half a year to an infinite number of times). Furthermore, the insurance fees corresponding to the continuous withdrawal case are very close to those corresponding to the half a year withdrawal case (the difference is less than 6 basis points for $\sigma = .2$ for the fourth refinement level).

In Figure 8.1, we show the value of the GMWB guarantee as a function of W at $t = 0$,

Refinement level	$\sigma = .20$	$\sigma = .30$
0	.0152023	.0317364
1	.0145009	.0313861
2	.0141471	.0312579
3	.0139699	.0312536
4	.0138905	.0312584
Value from [31]	.0137	.0304

TABLE 8.4: *Convergence study for the value of the fair insurance fee α, with respect to different values of σ. Data are given in Table 8.1. The value of α is computed so that the option value V satisfies $V = w_0 = 100$ at $t = 0$. Continuous withdrawal is permitted.*

Refinement	$\sigma = .20$		$\sigma = .30$	
level	$\Delta t_O = 1.0$	$\Delta t_O = .50$	$\Delta t_O = 1.0$	$\Delta t_O = .50$
0	.0128893	.0135554	.0291106	.0301345
1	.0128631	.0133379	.0292137	.0301367
2	.0128881	.0133312	.0292781	.0301912
3	.0129025	.0133441	.0293104	.0302238
4	.0129102	.0133516	.0293270	.0302407

TABLE 8.5: *Convergence study for the value of the fair insurance fee α in the discrete withdrawal case. Different withdrawal intervals Δt_O and different values of σ are considered. Data are given in Table 8.1. The value of α is computed so that the option value V satisfies $V = w_0 = 100$ at $t = 0$.*

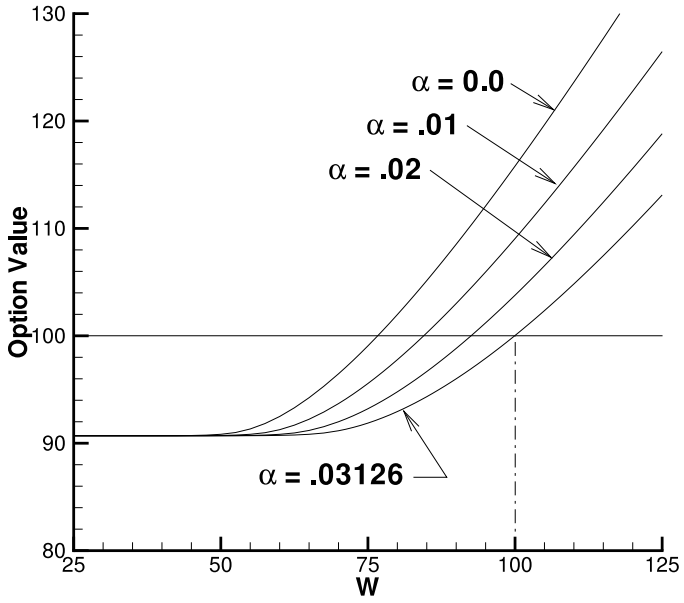


FIGURE 8.1: *The value of the GMWB guarantee as a function of W at $t = 0$, $A = 100$, with respect to various values of the insurance fee α corresponding to $W = 100$ including the fair value $\alpha = .03126$. The fair value of the fee occurs when the value of the guarantee V satisfies $V = w_0 = 100$. Data for this example are given in Table 8.1 with $\sigma = .30$. Continuous withdrawal is allowed.*

$A = 100$, with respect to various values of the insurance fee α including the fair value $\alpha = .03126$. The figure indicates that when W is relatively small, α has no effect on the contract value since in this case, the guarantee component of the contract dominates the equity component (i.e., $A \gg W$). Hence the contract value is determined only by the guarantee account value and is independent of the insurance fee which is imposed on the equity component. As the fee increases, the no-arbitrage value of the contract decreases near $W = 100$. Eventually, the value of the contract is precisely $V = 100$ at $W = 100$ when the fair fee is charged.

Figure 8.2 plots the value surface of the GMWB guarantee at $t = 0$ as a function of W and A assuming a fair insurance fee is imposed. The figure shows that the contract value increases as W and A increase. The value curve along the W direction transforms from a parabolic shape to a straight line as A changes from $A = 100$ to $A = 0$. Note that the surface forms a cusp along the line $A = W$ near $A = W = 0$.

We next study the optimal withdrawal strategy for an investor who maximizes the no-arbitrage value of the GMWB guarantee. More precisely, this is the worst case for

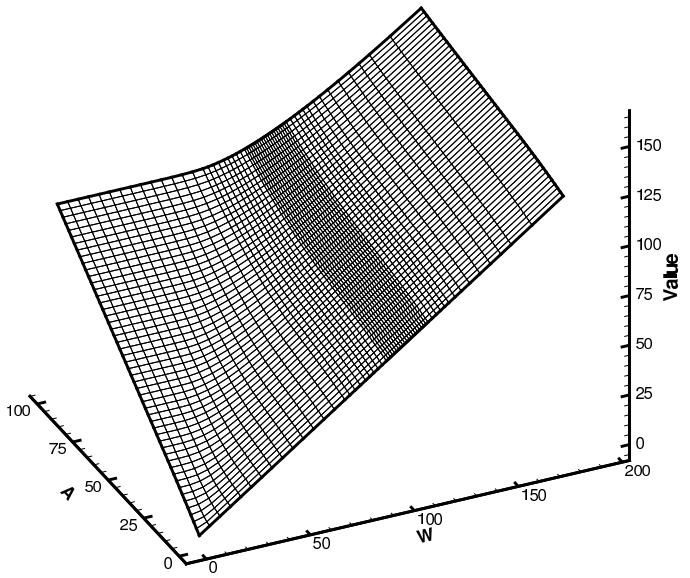


FIGURE 8.2: *The value of the GMWB guarantee at $t = 0$ as a function of sub-account balance W and guarantee account balance A. Data for this example are given in Table 8.1 with $\sigma = .30$ and the fair insurance fee $\alpha = .03126$. Continuous withdrawal is allowed.*

the provider of the guarantee. According to [31], the optimal strategy is either not to withdraw, or withdraw at the maximum rate G_r, or withdraw a finite amount instantaneously. Figure 8.3 shows a contour plot of the optimal withdrawal strategy at $t = \Delta\tau$ for different values of W and A computed numerically using the data from Table 8.1 with $\sigma = .3$ and using the fair insurance fee. From the figure, the (W, A)-plane is divided into a blank region and three shaded regions. The blank region corresponds to withdrawing continuously at the rate G_r. The upper left and upper right shaded areas correspond to withdrawing a finite amount instantaneously.

Within the elliptical shaded area in the lower left corner, our numerical results suggest zero withdrawals as the optimal strategy. This is unexpected. As is conjectured in [31], based on financial reasoning and numerical tests, it is never optimal not to withdraw since the investor will lose the proportional insurance fee α. To study the control behaviour within this region more carefully, we compute the ratio

$$R_{i,j} = \frac{V^h(W_i, A_j, \Delta\tau) - \left[V^h(W_i - G_r\Delta\tau, A_j - G_r\Delta\tau, \Delta\tau) + G_r\Delta\tau\right]}{G_r\Delta\tau}, \tag{8.7.1}$$

where $V^h(W_i, A_j, \Delta\tau)$ represents the approximate solution at the mesh node $(W, A, t) =$

$(W_i, A_j, \Delta\tau)$ and $V^h(W_i - G_r\Delta\tau, A_j - G_r\Delta\tau, \Delta\tau)$ is the corresponding approximate contract value after a withdrawal of $G_r\Delta\tau$. According to the optimization problem (7.4.2), if $R_{i,j} > 0$, our numerical scheme chooses a zero control at (W_i, A_j). If $R_{i,j} < 0$, the scheme suggests that it is optimal to withdraw at the rate G_r. We observe that for nodes residing within the shaded elliptical region, the ratios $R_{i,j}$ are positive but decrease towards zero quickly as we refine the mesh size (for example, the ratios are approximately 10^{-3} for the third refinement level). On the one hand, since the value of $|R_{i,j}|$ is insignificant, it is difficult for a numerical scheme to compute the sign of its exact value as $\Delta\tau \to 0$ due to numerical errors. As a result, $R_{i,j}$ may not have the same sign as its exact value and hence the zero withdrawal strategy returned by our scheme may not be correct. On the other hand, since the value of $|R_{i,j}|$ is very small, choosing $\hat{\gamma} = 0$ or $\hat{\gamma} = G_r$ will not affect the value of the guarantee. To verify this, we repeated the computation, but this time, we constrained the mesh nodes within the continuous withdrawal region to use the control value G_r, and disallowed zero as a possible control. The solution at $(W, A, t) = (100, 100, 0)$ resulting from this constraint is identical to the solution without imposing this constraint up to four digits.

To see this more clearly, assuming V^h is smooth and $\Delta\tau$ is sufficiently small, using Taylor series expansion leads to the approximation

$$R_{i,j} \approx V_W(W_i, A_j, \Delta\tau) + V_A(W_i, A_j, \Delta\tau) - 1. \tag{8.7.2}$$

Since we observe that $R_{i,j}$ converges to zero within this region, $V_W + V_A - 1$ also converges to zero in this region. According to the pricing equation (8.2.18), in this region, the equality

$$V_\tau - \mathcal{L}V - \sup_{\hat{\gamma}\in[0,G_r]} \left[\hat{\gamma}(1 - V_W - V_A)\right] = 0 \tag{8.7.3}$$

holds since the continuous withdrawal strategy is used. This implies that when $V_W + V_A - 1 \sim 0$, the optimal control $\hat{\gamma}$ can take any value between 0 and G_r, and hence is not unique. From the discussions above, our numerical results seem to suggest that $V_W + V_A - 1 = 0$ for mesh nodes within this region and thus the corresponding optimal

control is indeterminate, that is, any $\hat{\gamma} \in [0, G_r]$ is optimal.

The region of withdrawing a finite amount in the upper left of Figure 8.3 is also observed in [31]. In this region, W is less than A before the withdrawal; after the withdrawal, W decreases to zero and the investor carries on withdrawing the remaining balance from the guarantee account at the rate G_r. The strategy can be explained as follows. In this region, the guarantee account balance of the contract dominates the sub-account balance. Hence it is highly probable that the guarantee account value still dominates the sub-account balance, i.e., $A \gg W$, at maturity, and in this case the investor receives $(1-\kappa)A - c$ as the final payoff. In other words, the equity component has a small chance of contributing to the final payoff, but instead requires insurance fee payments. Consequently, it is optimal for the investor to withdraw all the funds from the sub-account (even subject to a penalty).

In Figure 8.3, the upper right region represents withdrawing a finite amount when the sub-account value W dominates the guarantee account value A. In this case, a finite withdrawal is optimal in order to reduce the insurance fee payment, since the guarantee has little value. Note that after the withdrawal, the sub-account balance still dominates the guarantee account value and can contribute to the contract payoff.

In the blank region of Figure 8.3, it is optimal to withdrawal at the rate G_r because this avoids the excessive withdrawal penalty due to withdrawing a finite amount and also avoids the additional insurance fee payment which would result if no withdrawals occurred.

8.8 Summary

Our work in this chapter is summarized as follows:

- We formulate the valuation of the GMWB variable annuities as an impulse control problem. As discussed in Remark 8.1, the impulse control formulation is a more general approach compared to the singular control formulation used in [69, 31].

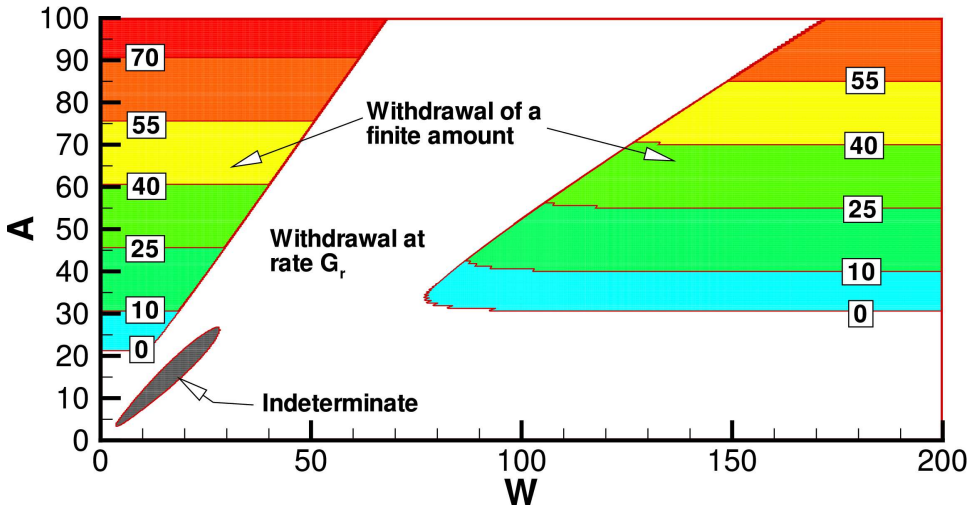


FIGURE 8.3: *The contour plot for the optimal withdrawal strategy of the GMWB guarantee at $t = \Delta\tau$ in the (W, A)-plane. In the regions of withdrawing finite amounts, contour lines representing the same withdrawal levels are shown, where the withdrawal amounts are posted on those contour lines. Data for this example are given in Table 8.1 with $\sigma = .30$ and the fair insurance fee $\alpha = .03216$. Continuous withdrawal is allowed. In the region labeled* indeterminate*, the numerical results indicate that the same value is obtained for any control rate in $[0, G_r]$.*

- We generalize the scheme for the discrete withdrawal case to the continuous withdrawal case. As shown in Remark 8.2, the scheme is identical to an extension of the semi-Lagrangian timestepping method for the bounded stochastic control problems, introduced in Chapter 2. Therefore, we have a single numerical method for solving the bounded and unbounded stochastic control problems as well as their discrete versions where the operations are performed only at discrete times.

- Provided a strong comparison result holds, we prove that the scheme converges to the unique viscosity solution of the HJB variational inequality corresponding to the impulse control problem by verifying the l_∞-stability, monotonicity and consistency of the scheme and using the basic results in [11, 6].

- We provide some numerical tests which indicate that the no-arbitrage fee for the discrete withdrawal contract is very close to the continuous contract fee (i.e. to within a few basis points) even for fairly infrequent withdrawal intervals (e.g. once every half a year).

- For the continuous withdrawal case, the numerical results suggest that our impulse control formulation converges to the singular control formulation in [31] as the fixed cost vanishes. The numerical results also demonstrate that our scheme can solve the impulse control problem with a nonzero fixed cost as well as the singular control problem by setting the fixed cost to be zero, although the convergence is proved only for the former case.

- Our numerical results appear to show that the optimal control strategy may not be unique. That is, there exists a region where different control strategies can result in the same guarantee value.

Chapter 9

The Effect of Modelling Parameters on the Value of GMWB Guarantees

In previous chapters, we have introduced GMWB variable annuities and proposed numerical schemes for pricing the contracts. In this chapter, we conduct an extensive study of the no-arbitrage fee for GMWB contacts assuming various parameters and contract details.

In particular, we study the following:

- In practice, the underlying mutual fund charges a separate layer of fees for managing the fund. It has been suggested that the apparent underfunding of the GMWB rider can be explained if we assume that some of these mutual fund fees are diverted to managing the GMWB rider. However, as suggested in [90, 88, 63], the mutual fund fees are often not available for hedging purposes. We will derive the no-arbitrage PDE which results from this fee splitting, and provide numerical results which show the effect of the fee splitting. This fee separation is important in practice, and does not appear to have been taken into account in previous work (e.g., [69, 31, 13]). Inclusion of this fee separation increases the value of the GMWB rider.

- The authors of [28] discuss various assumptions about investor behaviour when pricing variable annuities. A conservative approach is to assume optimal investor

behaviour, and then to recognize extraordinary earnings in the event of sub-optimal behaviour. Another possibility is to develop a model of non-optimal behaviour, and incorporate this into the pricing model. We will examine both approaches in this chapter. Our base case assumes optimal behaviour, but we also model the effect of sub-optimal withdrawal behaviour using the approach suggested in [54]. Sub-optimal behaviour considerably reduces the value of the GMWB rider.

- We will include results with both the classic Geometric Brownian motion process for the underlying asset, as well as a jump diffusion process, which may be a more realistic model for long term guarantees. Making the assumption that there is reasonable (risk neutral) probability of a market crash during the lifetime of the guarantee dramatically increases the value of the GMWB rider.

- We will also examine the effect of various contract parameters, such as reset provisions, maturities, withdrawal intervals, and surrender charges. Some contract features (e.g. the reset provision) have almost no effect on the value of the guarantee, while others have considerable influence. In addition to the value of the guarantee, we explore the impact of some of these various alternative modelling assumptions on the policyholder's optimal withdrawal strategy. Plots of the contour levels of the optimal withdrawal amounts show that the investor's optimal strategy can be quite sensitive to modelling assumptions.

9.1 The Mathematical Model

We assume that the withdrawal occurs only at predetermined discrete times (see Chapter 7 for the description of a simple discrete withdrawal contract). This problem can also be posed in terms of continuous withdrawals described in Chapter 8. However, as we verify through numerical experiments in Chapter 8 and this chapter, the value of the continuous withdrawal formulation is very close to the discrete withdrawal case if the withdrawal intervals are less than one year. In addition, many contracts only allow discrete withdrawals.

Year	Surrender Charge κ
$0 \le t < 2$	8%
$2 \le t < 3$	7%
$3 \le t < 4$	6%
$4 \le t < 5$	5%
$5 \le t < 6$	4%
$6 \le t < 7$	3%
$t \ge 7$	0%

TABLE 9.1: *Time-dependent surrender charges $\kappa(t)$.*

The problem notation has been given in Chapter 7. In contrast to the contract described in Chapter 7, in this chapter we assume the guarantee fees include a mutual fund management fee. Let $\alpha_{tot} \ge 0$ denote the proportional total fees (including mutual fund management expenses and fees charged for the GMWB rider) paid by the policy holder. Let α_g be the fee paid to fund the guarantee, and α_m be the mutual fund management fee, so that $\alpha_{tot} = \alpha_g + \alpha_m$. Given this notation, according to (7.2.3-7.2.4), the risk neutral process of the sub-account value W is given by

$$dW = (r - \alpha_{tot})W dt + \sigma W dZ, \quad \text{if } W > 0 \tag{9.1.1}$$

$$dW = 0, \quad \text{if } W = 0. \tag{9.1.2}$$

In a typical contract, the deferred surrender charge $\kappa = \kappa(t)$ is time-dependent and normally decreases over time to zero. Table 9.1 shows a typical specification for $\kappa(t)$. The cash flow received by the investor after a withdrawal of γ^k at the withdrawal time τ_O^k is given in (7.2.5).

The terminal condition for the annuity is given in (7.2.9).

At the withdrawal time $\tau = \tau_O^k$, V satisfies the optimality condition (7.2.6).

Within each time interval $[\tau_O^{k+}, \tau_O^{k+1}]$, $k = 0, \ldots, K-1$, the annuity value function $V(W, A, \tau)$, solves the following linear PDE which has A dependence only through equation (7.2.6):

$$V_\tau = \mathcal{L}V + \alpha_m W, \quad \tau \in [\tau_O^{k+}, \tau_O^{k+1}], \quad k = 0, \ldots, K-1. \tag{9.1.3}$$

	Parameter	Value
T	Expiry time	10 years
r	Interest rate	5%
G	Contract withdrawal amount	10
w_0	Initial lump-sum premium	100
σ	Volatility	.15
Δt_O	Withdrawal interval	1 year
α_m	Mutual fund fee	1%

TABLE 9.2: *Base case parameters.*

where the operator $\mathcal{L}$ is

$$\mathcal{L}V = \frac{1}{2}\sigma^2 W^2 V_{WW} + (r - \alpha_{tot})WV_W - rV. \tag{9.1.4}$$

The derivation of (9.1.3) is given in Appendix I, based on a no-arbitrage hedging argument.

The fair insurance fee α_g (recall $\alpha_{tot} = \alpha_g + \alpha_m$) is determined so that the contract value V at $\tau = T$ is equal to the initial premium w_0 paid by the investor [69, 31, 28]. The pricing equations are solved numerically using the methods described in Chapter 7. In the following, all results are given correct to the number of digits shown, based on grid and timestep refinement studies.

9.2 Numerical Results

9.2.1 Base Case

We first compute the value of the guarantee for a representative base case, and then perturb the problem parameters and compare to this base case. The base case parameters are given in Table 9.2, with the surrender charge $\kappa(t)$ (see equation (7.2.5)) given in Table 9.1.

For this base case, the no-arbitrage insurance fee is $\alpha_g = 117$ basis points. Figure 9.1 shows a contour plot of the optimal withdrawal strategy γ^k at the first withdrawal time ($t = 1$) for different values of W and A. In particular, we show contour levels of $\gamma =$

$10, \gamma = .2$. We choose these two contour levels to show that in some cases the optimal withdrawal amount γ rapidly changes from the contract amount G^k to zero. (Due to contouring artifacts, a contour value $\gamma < .2$ results in very jagged contour levels, since it is difficult to determine numerically the zero withdrawal region). For practical purposes, the $\gamma = .2$ contour level shows the region where it is optimal to withdraw nothing. For a discussion of the conditions under which it may be optimal to withdraw nothing, see Chapter 8.

From Figure 9.1, we can observe the following:

- There is a shaded region in the left side of the figure representing excessive withdrawals (i.e. withdrawals above the contract amount) when A dominates W. In this region, it is unlikely that the amount in the risky sub-account will ever exceed the guarantee account. Intuitively, the investor withdraws as rapidly as possible (subject to minimizing the surrender charges) because the total guarantee available is just w_0 and delaying withdrawal is costly due to a lower present value of the funds withdrawn. As a specific example, consider the case where $(W, A) = (0.0, 80)$. From Table 9.1, the investor will receive .92 of any withdrawal above 10. Also, note (from Table 9.2) that $r = .05$. In this case, it is optimal to withdraw 70 immediately, and then withdraw 10 the next year. The present value of this strategy is $10 + 60 \times .92 + e^{-.05}10 = 74.71$. This is slightly better than withdrawing 80 immediately, which has value of $10 + 70 \times .92 = 74.40$.

- There is a blank region in the right side of the figure representing the withdrawal of the contract amount $\gamma^k = G^k$.

- There is a narrow area surrounding the line $W = A$, in which the optimal strategy is to withdraw an amount less than G^k. The contour line for withdrawing $\gamma^k = .2$ is also shown, to illustrate the rapid change in the withdrawal amount over a small region in the (W, A) plane.

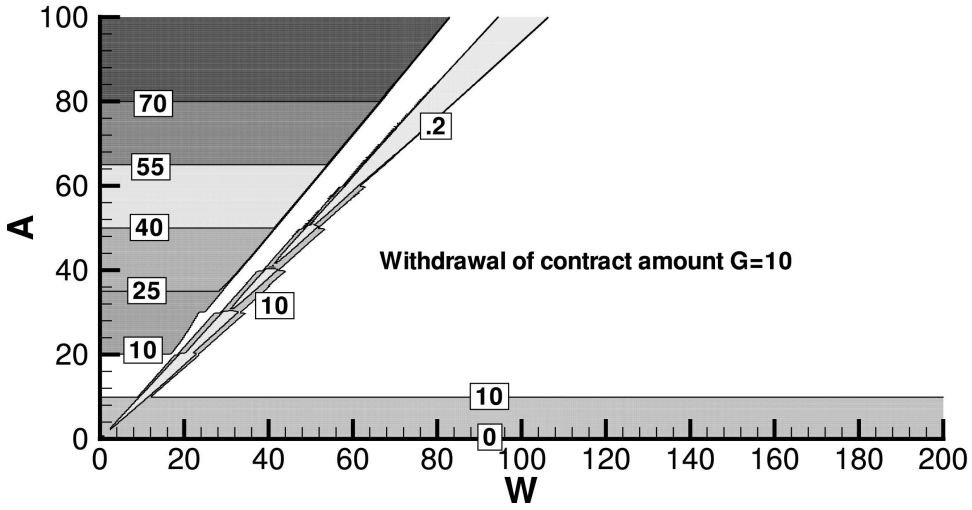


FIGURE 9.1: *Optimal withdrawal strategy of the GMWB guarantee at the first withdrawal time ($t = 1$ year) in the (W, A)-plane. Contour lines representing the same withdrawal levels are shown, where the withdrawal amounts are posted on those contour lines. Parameters for this example are given in Tables 9.1 and 9.2. The contour line showing $\gamma = .2$ shows the region where it is optimal to withdraw essentially nothing. Fair insurance fee $\alpha_g = 117$b.p.*

Volatility σ	Insurance Fee α_g
.15	117 b.p.
.20	214 b.p.
.25	326 b.p.
.30	440 b.p.
.35	552 b.p.

TABLE 9.3: *GMWB guarantee fees α_g determined with different choices of the volatility σ. Other parameter values are given in Tables 9.1 and 9.2.*

9.2.2 Effect of Volatility

The insurance fees for different choices of volatility σ are given in Table 9.3. The table shows that volatility has a large effect on the no-arbitrage value of the guarantee fee α_g. For example, the fee level almost doubles when σ is set to .20 compared to the base case of $\sigma = .15$.

Figures 9.2, 9.3 and 9.4 show the optimal withdrawal strategy for $\sigma = .20$ at the first, fourth and eighth withdrawal time forwards in time (with respect to $t = 1$st year, $t = 4$th year and $t = 8$th year). The figures reveal that:

- Compared with Figure 9.1, increasing volatility generates another excessive withdrawal region when W dominates A. When $W \gg A$, the guarantee is effectively out the money. Hence the investor withdraws an amount which minimizes the fees charged for this out of the money guarantee, subject to minimizing the surrender charges.

- As t increases, the shaded regions representing excessive withdrawals expand. At the same time, the blank region representing the withdrawal of the contract amount G^k shrinks. This is due to the decrease of the surrender charge κ over time, which imposes a smaller penalty on excessive withdrawals. However, it is also interesting to note that the no-withdrawal region (i.e. the region enclosed by the $\gamma = .2$ contour) expands as well.

- Excessive withdrawals at a later time (i.e. a larger t) will result in an equal or less remaining balance in the guarantee account compared with excessive withdrawals at an earlier time (i.e. a smaller t). In particular, at the eighth withdrawal time when $\kappa = 0$, the optimal strategy is to withdraw the whole amount from the guarantee account in the left shaded region (where A dominates W) as well as in the right shaded region (where W dominates A), excluding the triangular area surrounding the line $W = A$. The area near $W = A$ can be regarded as an at the money put option. Since there are only two years left in the contract, the fees charged for this at the money put are comparatively low, and hence it is worthwhile for the policyholder to keep the option intact (i.e. not to withdraw).

9.2.3 Incorporating Price Jumps

Many studies have shown that for long term contingent claims, it is important to consider jump processes [4]. For this example, we assume that the dynamics of W follows a jump

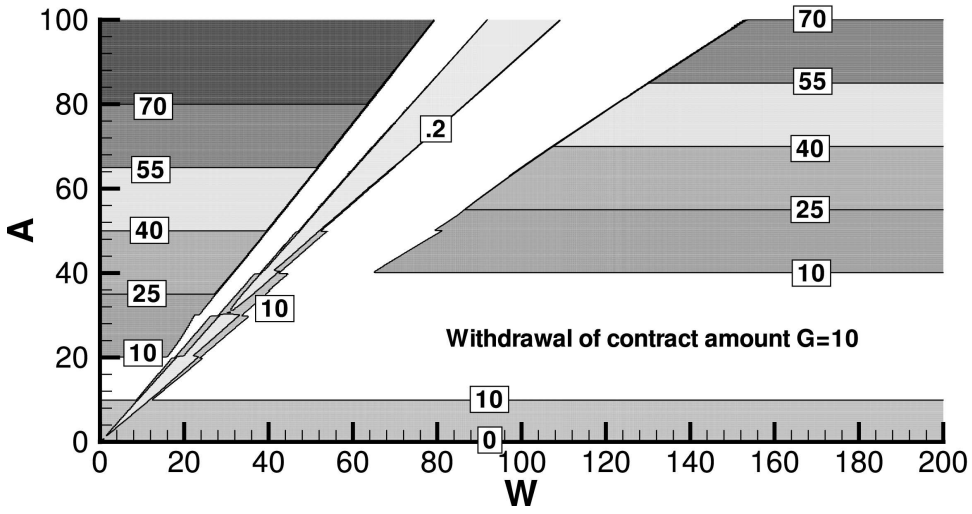


FIGURE 9.2: *Optimal withdrawal strategy of the GMWB guarantee at the first withdrawal time forwards in time ($t = 1$ year) in the (W, A)-plane with $\sigma = .20$. Other parameter values are given in Tables 9.1 and 9.2. The contour line showing $\gamma = .2$ shows the region where it is optimal to withdraw essentially nothing.*

diffusion process given by

$$dW = (r - \alpha_{tot} - \lambda\beta)W dt + \sigma W dZ + (\eta - 1)W dq, \quad \text{if } W > 0 \tag{9.2.1}$$

$$dW = 0, \quad \text{if } W = 0, \tag{9.2.2}$$

where:

- dq is an independent Poisson process with $dq = \begin{cases} 0 & \text{with probability } 1 - \lambda dt \\ 1 & \text{with probability } \lambda dt \end{cases}$,
- λ is the jump intensity representing the mean arrival rate of the Poisson process,
- η is a random variable representing the jump size of W; we assume that η follows a log-normal distribution $g(\eta)$ given by

$$g(\eta) = \frac{1}{\sqrt{2\pi}\zeta\eta} \exp\Big(-\frac{(\log(\eta) - \nu)^2}{2\zeta^2}\Big) \tag{9.2.3}$$

  with parameters ζ and ν,
- $\beta = E[\eta - 1]$, where $E[\eta] = \exp(\nu + \zeta^2/2)$ given the distribution function $g(\eta)$ in

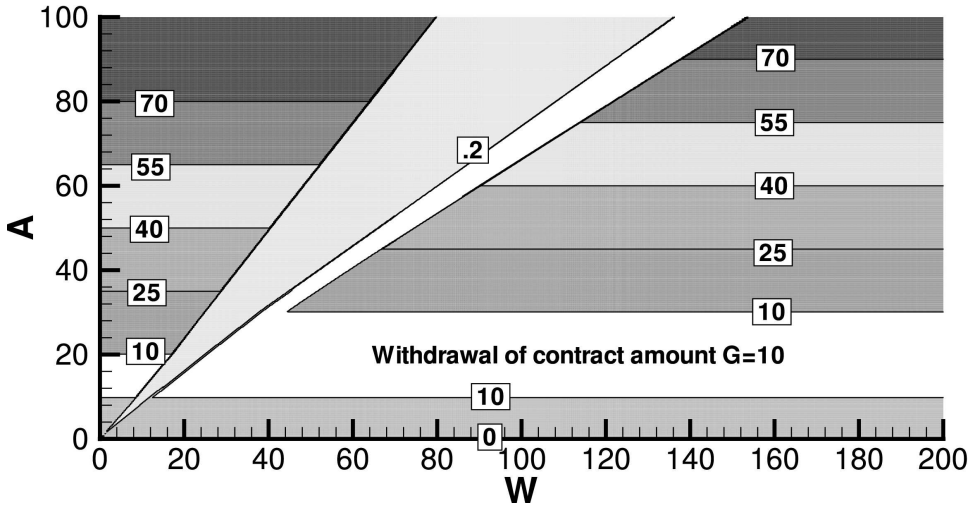


FIGURE 9.3: *Optimal withdrawal strategy of the GMWB guarantee at the fourth withdrawal time forwards in time ($t = 4$ years) in the (W, A)-plane with $\sigma = .20$. Other parameter values are given in Tables 9.1 and 9.2. The contour line showing $\gamma = .2$ shows the region where it is optimal to withdraw essentially nothing.*

(9.2.3).

Note that we are working here in an incomplete market, so that the equivalent martingale pricing measure is not in general unique. As in [4], we can calibrate the parameters of equation (9.2.1) to traded prices of options. This means that the parameters of (9.2.1) will correspond to those from the market's pricing measure.

Using (9.2.1), it is straightforward to to generalize the pricing PDE (9.1.3) to the pricing partial integrodifferential equation (PIDE)

$$V_\tau - \mathcal{L}V - \mathcal{H}V - \alpha_m W = 0, \quad \tau \in [\tau_O^{k+}, \tau_O^{k+1}], \quad k = 0, \ldots, K-1, \tag{9.2.4}$$

where the operator $\mathcal{H}$ satisfies

$$\begin{aligned} \mathcal{H}V &= \lambda E\big[V(W\eta) - V - (\eta - 1)WV_W\big] \\ &= \lambda \int_0^\infty V(W\eta)g(\eta)d\eta - \lambda V - \lambda\beta W V_W. \end{aligned} \tag{9.2.5}$$

We solve PIDE (9.2.4) using the numerical methods described in [41]. Table 9.4 gives the jump diffusion parameters which we will use in our tests. These parameters are

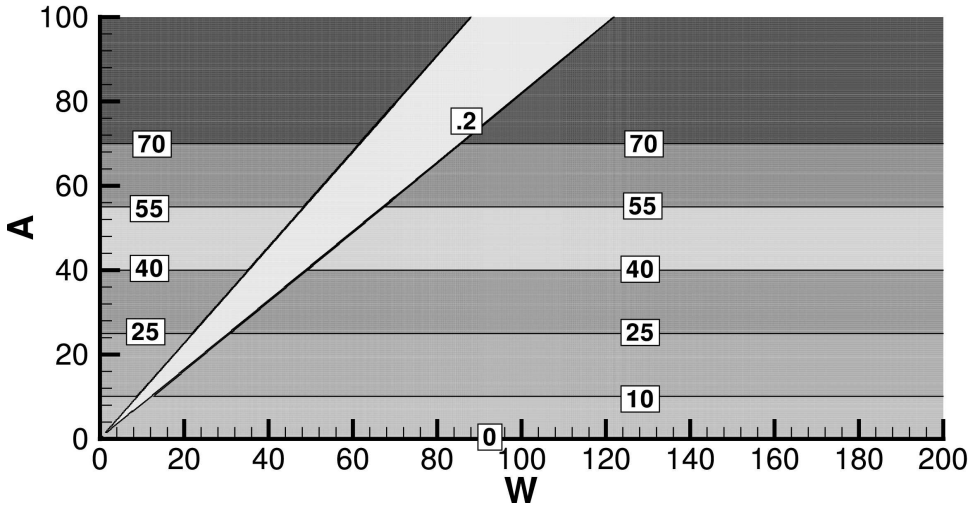


FIGURE 9.4: *Optimal withdrawal strategy of the GMWB guarantee at the eighth withdrawal time forwards in time ($t = 8$ years) in the (W, A)-plane with $\sigma = .20$. Other parameter values are given in Tables 9.1 and 9.2. The contour line showing $\gamma = .2$ shows the region where it is optimal to withdraw essentially nothing.*

Parameter	Value
λ	.1
ζ	.45
ν	-.9
σ	.15

TABLE 9.4: *Parameters for the jump diffusion case.*

essentially the (rounded) parameters obtained in [4] from calibration to S&P 500 index option prices.

The fair insurance fees for the jump and no jump cases are given in Table 9.5. The table shows that incorporating jumps greatly increases the insurance fees. Note that a volatility of $\sigma = .15$ may appear to be reasonable if one examines recent long term data for implied volatility of a major stock index. However, the implied volatilities are based on short term options, which do not capture long term information about jumps. Table 9.5 shows that ignoring the possibility of jumps for long term guarantees may severely underestimate the hedging cost. Methods for hedging contracts under jump diffusions are discussed in [53, 58].

	Insurance Fee α_g
With jump	356 b.p.
Base case (no jump)	117 b.p.

TABLE 9.5: *Insurance fees with/without price jumps. Parameters are given in Tables 9.1, 9.2 and 9.4.*

Mutual Fund Fee α_m	Insurance Fee α_g
0.0%	88 b.p.
0.5%	102 b.p.
1.0%	117 b.p.
1.5%	136 b.p.
2.0%	157 b.p.
2.5%	184 b.p.

TABLE 9.6: *Insurance fees determined by different choices of the mutual fund fees. Other parameter values are given in Tables 9.1 and 9.2.*

9.2.4 Separation of Mutual Fund Fee

Recall that in equation (9.1.3), we are careful to distinguish between the fees used to manage the underlying mutual fund (α_m) and the fees used to fund the GMWB guarantee (α_g). This is because the provider of the GWMB rider may be a completely separate business unit from the unit managing the mutual fund [63]. A precise hedging scenario for this case is given in Appendix I.

The no-arbitrage GMWB guarantee fees α_g for different choices of the mutual fund fees α_m (assuming no jumps in W) are given in Table 9.6. The table shows that the GMWB fee α_g is very sensitive to the mutual fund fee α_m. The effect of the mutual fund fee on the GMWB fee has not been taken into account previously. Note that a mutual fund fee of $\alpha_m = 1.0\%$ increases the GMWB fee by 29 b.p. compared to the case where $\alpha_m = 0$. The intuition for this is straightforward: the guarantee applies to the initial value of the account w_0, prior to any fees being deducted. As the mutual fund fees increase, the account value is correspondingly reduced over time, thereby increasing the value of the guarantee.

	Insurance Fee α_g
Constant κ	95 b.p.
Decreasing κ	117 b.p.

TABLE 9.7: *Insurance fees for constant/decreasing κ. For the decreasing $\kappa(t)$ case, the data is given in Table 9.1. For the constant κ case, the flat rate is 8%. Other parameter values are in Table 9.2.*

9.2.5 Constant Surrender Charge

Table 9.7 provides the GMWB guarantee fees for a constant surrender charge $\kappa = 8\%$ and for the base case $\kappa(t)$ as in Table 9.1. It is perhaps surprising that the GMWB fee does not appear to decrease greatly for the case where the surrender fee is constant compared to the case where $\kappa(t)$ decreases to zero. Intuitively, one reason for this is because the reported values are at $t = 0$, and the benefit to the investor of the reduced surrender charge is discounted for a relatively lengthy period of time. Moreover, this benefit only arises in states in which it is optimal to withdraw an amount greater than that contractually specified. It is also worth recalling that we are assuming that the surrender charges can be used to fund the guarantee. In general, though, decreasing the fee to zero appears to be mainly a marketing tool, rather than a valuable option for the investor.

9.2.6 Sub-optimal Control Strategy

The previous results were computed assuming an optimal withdrawal policy by the GMWB contract holder. The issue of how to model consumer behaviour when pricing and hedging variable annuities is controversial. It is instructive to reproduce a quote from [28]:

> *"Assumptions should reflect that an option will impact policyholder behaviour, and the degree to which it impacts policyholder behaviour will be a function of how much the option is in the money...*
>
> *Some actuaries believe that all policyholders should be expected to always act optimally, and earnings only recognized when sub-optimal behaviour occurs.*

Because the valuation is typically done using risk-neutral assumed returns, some actuaries believe it is appropriate to adjust policyholder behaviour assumptions to reflect policyholder decisions based on a 'real world' environment. Others believe that this approach is inconsistent with a risk neutral framework."

We model non-optimal behaviour using the method suggested in [54]. We consider here the sub-optimal behaviour of the investor described as follows: at each withdrawal time τ_O^k, the default strategy of the investor is to precisely withdraw the contract withdrawal amount G^k. Nevertheless, the investor will switch to the optimal withdrawal strategy (if it is different from the default strategy) if the difference between the value corresponding to the optimal strategy and the value corresponding to the default strategy is no less than a fraction S of the initial lump-sum payment. This makes it more likely that the holder will act optimally if the option is deep in the money. To be more precise, the value $V(\tau_O^{k+})$ instantaneously following each withdrawal time τ_O^k is determined by Algorithm 9.1.

// Value obtained from the default strategy
$U_d(W, A, \tau_O^{k+}) = V\big(\max(W - \gamma_*^k, 0), A - \gamma_*^k, \tau_O^k\big) + f(\gamma_*^k)$, where $\gamma_*^k = \min(A, G^k)$

// Value obtained from the optimal strategy
$U_o(W, A, \tau_O^{k+}) = \sup_{\gamma^k \in [0,A]} \big[V\big(\max(W - \gamma^k, 0), A - \gamma^k, \tau_O^k\big) + f(\gamma^k)\big]$

// Model the sub-optimal behaviour by comparing the above two values
If $U_o(W, A, \tau_O^{k+}) - U_d(W, A, \tau_O^{k+}) \geq Sw_0$ Then
  $V(W, A, \tau_O^{k+}) = U_o(W, A, \tau_O^{k+})$
Else
  $V(W, A, \tau_O^{k+}) = U_d(W, A, \tau_O^{k+})$

ALGORITHM 9.1: *Sub-optimal behaviour*

Effectively, we are assuming that the holder will not bother to withdraw optimally, unless the optimal strategy is considerably more valuable than the default strategy (scaled by the initial lump-sum payment). Table 9.8 gives the fair insurance fees obtained under

Parameter S	$\sigma = .15$	$\sigma = .20$
0 (Always withdraw optimally)	117 b.p.	214 b.p.
0.03	86 b.p.	162 b.p.
0.05	77 b.p.	150 b.p.
∞ (Always withdraw G^k)	64 b.p.	123 b.p.

TABLE 9.8: *Insurance fees for the sub-optimal and optimal control strategies. Other parameter values are given in Tables 9.1 and 9.2. It is assumed that the holder will take the default action (withdraw at contract rate) unless the gain in acting optimally is larger than Sw_0, where w_0 is the initial lump-sum payment. See Algorithm 9.1 for details.*

the above sub-optimal behaviour for different choices of the volatility σ and the optimality parameter S. We can see from Table 9.8 that if the holder always withdraws at the default rate, then the value of the guarantee is reduced by about one-half. However, the sub-optimal GMWB guarantee is still quite valuable for moderate volatilities (123 b.p. for $\sigma = .20$).

9.2.7 Reset Provision

We next consider a reset provision in the contract. This feature was discussed in [69, 31], but no results were reported for this case in those papers. Under this provision, if an excessive withdrawal occurs at some withdrawal time, that is, if $\gamma^k > G^k$ for some k, then the guarantee account balance A is reset to the minimum of $A - \gamma^k$ and the resulting sub-account value, $\max(W - \gamma^k, 0)$. Let $\mathcal{A}^k$ denote the remaining guarantee account balance after withdrawal. Then we have

$$\mathcal{A}^k = \begin{cases} A - \gamma^k & \text{if } \gamma^k \le G^k \\ \min\{A - \gamma^k, \max(W - \gamma^k, 0)\} & \text{if } \gamma^k > G^k \end{cases}. \tag{9.2.6}$$

In this case, the optimality condition (7.2.6) becomes

$$V(W, A, \tau_O^{k+}) = \sup_{\gamma^k \in [0, A]} \left[V\big(\max(W - \gamma^k, 0), \mathcal{A}^k, \tau_O^k\big) + f(\gamma^k) \right], \quad k = 0, \ldots, K-1. \tag{9.2.7}$$

	$\sigma = .15$	$\sigma = .20$
With reset provision	116 b.p.	212 b.p.
Without reset provision	117 b.p.	214 b.p.

TABLE 9.9: *Insurance fees with/without reset provision. Other parameter values are given in Tables 9.1 and 9.2.*

Table 9.9 provides the fair insurance fees after imposing the reset provision for different choices of volatility. The table shows that the reset provision has little effect on the insurance fee.

Figure 9.5 shows the optimal withdrawal strategy at the first withdrawal time after imposing reset provision, where we set $\sigma = .20$. Compared with the control strategy without imposing reset provision in Figure 9.2, we can observe that:

- The right side excessive withdrawal region in Figure 9.5 is identical to that in Figure 9.2.
- The left side excessive withdrawal region in Figure 9.5 amounts to withdrawing all the remaining guarantee account balance, as opposed to the corresponding region in Figure 9.2, where the remaining balance after excessive withdrawal is 10. This is due to the extra penalty imposed on A (see equation (9.2.6)) that promotes complete withdrawal in order to reduce the resulting loss. Even though the addition of the reset provision (9.2.6-9.2.7) results in a different optimal strategy when $A \ll W$ compared with the base case, these regions clearly must have a low (risk neutral) probability, so that the guarantee value is almost unchanged (see Table 9.9).

9.2.8 Different Maturities

Table 9.10 gives the fair insurance fees with respect to different maturities. The contract withdrawal rate $G^k = w_0/T$ for these cases. As might be expected, the value of the guarantee decreases as the maturity increases, due to the reduced time value of the guarantee (recall that the total guaranteed withdrawal amount initially is w_0, irrespective of the maturity date).

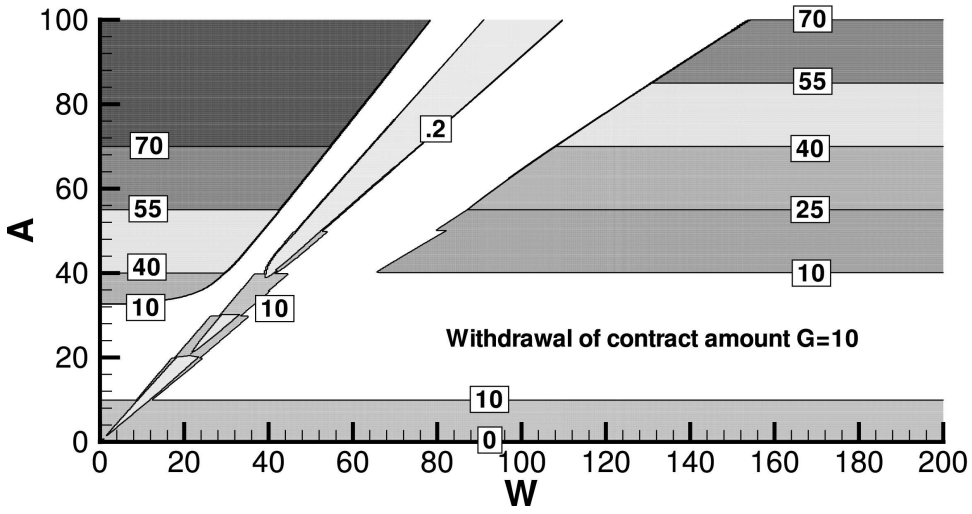


FIGURE 9.5: *Optimal withdrawal strategy of the GMWB guarantee at the first withdrawal time forwards in time ($t = 1$ year) in the (W, A)-plane with reset provision imposed and with $\sigma = .2$. Other parameters are given in Tables 9.1 and 9.2. The contour line showing $\gamma = .2$ shows the region where it is optimal to withdraw essentially nothing.*

Maturity T	Insurance Fee α_g
5 Years	183 b.p.
10 Years	117 b.p.
20 Years	79 b.p.

TABLE 9.10: *Insurance fees for different maturities. Other parameters are given in Tables 9.1 and 9.2. The contract withdrawal rate $G^k = w_0/T$ for these cases.*

9.2.9 Different Withdrawal Intervals

Table 9.11 gives the fair insurance fees with respect to different withdrawal intervals. In this case, the maximum withdrawal without penalty is given by $G^k = 10\Delta t_O$, where Δt_O is the withdrawal interval. The table shows that the withdrawal intervals have a fairly small effect on the insurance fees. For example, decreasing the withdrawal interval from one year to one month increases the fee by only 5 b.p. This means that allowing more frequent withdrawals (but at the same yearly rate) does not increase the value of the guarantee significantly.

Withdrawal Interval	Insurance Fee α_g
2 Years	107 b.p.
1 Year	117 b.p.
6 Months	119 b.p.
1 Month	122 b.p.

TABLE 9.11: *Fair insurance fees for different choices of withdrawal intervals. Other parameters are given in Tables 9.1 and 9.2.* $G^k = 10\Delta t_O$*, where* Δt_O *is the withdrawal interval.*

Interest Rate	Insurance Fee α_g
1%	761 b.p.
3%	227 b.p.
5%	117 b.p.
7%	68 b.p.
9%	41 b.p.

TABLE 9.12: *Fair insurance fees for different values of* r*. Other parameters are given in Tables 9.1 and 9.2.*

9.2.10 Varying Interest Rates

Table 9.12 gives the fair insurance fees with respect to different values of the risk free interest rate r. The guarantee values are extremely sensitive to interest rates, due to the time value of the guarantee account. A reduction in the risk free rate drastically increases the value of the GMWB guarantee.

9.3 Summary

Our work in this chapter is summarized as follows:

- We have carried out an extensive analysis of the no-arbitrage fee for GMWB guarantees. Typical fees for GMWB guarantees are less than 50 b.p. [69]. Fees of this level can be justified (assuming that the guarantee is hedged) only if all of the following assumptions are made:

  - Volatilities are 15% or less. Market crashes (jumps) are either impossible or at least extremely unlikely over the lifetime of these contracts (under the pricing measure).
  - Contract holders withdraw sub-optimally.
  - Interest rates do not drop to relatively low levels.

- In all other cases, the GMWB no-arbitrage fee should be considerably higher than 50 b.p. This would suggest that insurers are exposed to considerable risk, since the fees collected are not enough to cover hedging costs.

- In many cases, the business unit hedging the guarantee is separate from the unit managing the mutual fund. For typical parameters, the GMWB hedging fee increases by about 30 b.p. due to this fee splitting.

- We note that most GMWB contracts have additional optionality (e.g. ratchet provisions and increases in the guarantee account if withdrawals are skipped), which can only increase the value of the guarantee. Finally, we have implicitly assumed throughout this chapter (this is made explicit in Appendix I) that the insurer can hedge these contractual obligations using a traded instrument that is perfectly correlated with the investment in the investor's account. In other words, we have ignored basis risk. Of course, our stochastic modelling assumptions are also rather simple (e.g. constant volatility and interest rates). Hedging risk exposures arising from such real world features would require additional reserves. Consequently, the guarantee values calculated in this chapter will undervalue typical contracts. This reinforces our primary conclusion that insurance companies are not charging enough to hedge these contracts.

Chapter 10

Conclusion

In this thesis, we develop partial differential equation (PDE) based numerical methods to solve the optimal stochastic control problems in finance.

A stochastic control problem is linked with a nonlinear PDE in the viscosity sense, that is, the solution of the stochastic control problem is normally identical to the viscosity solution of the corresponding PDE. Therefore, the solution to the stochastic control problem can be computed by solving the corresponding PDE numerically, as long as the convergence to the viscosity solution is guaranteed.

Stochastic control problems with bounded controls correspond to Hamilton-Jacobi-Bellman (HJB) equations. We develop a fully implicit scheme based on a semi-Lagrangian timestepping to discretize the HJB equation. Initially introduced by [42, 74] for atmospheric and weather numerical predictions, semi-Lagrangian schemes can effectually reduce the numerical problems arising for convection dominated equations. Our numerical scheme has the following desired properties:

- The scheme is unconditionally stable, i.e., there are no timestep limitations due to stability considerations.
- The scheme can be shown rigorously to converge to the viscosity solution.
- The semi-Lagrangian timestepping completely separates the inventory component

from the underlying stochastic model. Thus, it can easily handle more complex stochastic models such as jump diffusion and regime-switching models.

- The semi-Lagrangian timestepping is more efficient than standard implicit finite difference discretizations given in [46], since it avoids Policy type iterations at each mesh node at each timestep. Instead, it requires solution of a local optimization problem at the mesh point, which can be solved efficiently. The semi-Lagrangian method also reduces the problem to a set of independent sub-problems, which makes it ready for parallel implementation.

- The scheme is algebraically identical to a discretization based on a scenario that the operations are performed only at discrete times.

We demonstrate the properties of our scheme by valuing natural gas storage facilities using a one-factor mean-reverting model for natural gas spot prices. We then extend the scheme to solve the same problem assuming a more complex regime-switching model for natural gas spot, which we demonstrate by calibration is able to fit the market data more accurately than the one-factor mean-reverting model.

We then study the stochastic control problems with unbounded controls. We model the unbounded control problems as impulse control problems, resulting in HJB variational inequalities. We show that the impulse control framework is more general than the singular control framework. We then generalize the semi-Lagrangian timestepping for the bounded control case to solve the HJB variable inequality corresponding to the impulse control formulation.

The generalized scheme also satisfies the desired properties above except that in the impulse control case the local optimization problem at each mesh point will be more costly to solve compared with the existing penalty methods [47, 31], although the penalty methods require nonlinear iterations. However, we have a unified numerical method for both the bounded and the unbounded control problems as well as the discrete cases where the controls are allowed only at discrete times.

We demonstrate the properties of the semi-Lagrangian timestepping for the unbounded

control case by pricing the variable annuities with guaranteed minimum withdrawal benefits (GMWBs).

10.1 Future Work

There are some research directions that can be pursued in the future work.

- The optimization problem in the impulse control case is expensive to solve since a search across the whole discretization grid is required. It would be desirable to develop a method that is able to solve the optimization problem more efficiently.
- The penalty method can be used to solve the singular control formulation; no optimization problem needs to be solved in this case. It would be interesting to compare our semi-Lagrangian discretization with the penalty method in terms of running time and accuracy.
- Finally, it would be interesting to apply our scheme to solve other stochastic control problems such as portfolio selection problems, optimal hedging etc.

Appendix A

Derivation of Natural Gas Storage Pricing Equation

This chapter heuristically derives the gas storage pricing PDE from the stochastic control formulation (2.2.8) using dynamic programming (Bellman's Principle) and Ito's Lemma. We assume $\hat{V}$ is smooth.

We first rewrite the equation (2.2.8) as

$$
\begin{aligned}
&\hat{V}(P,I,t) \\
&= \sup_{c(s)\in C(I(s))} E^{\mathbb{Q}}\Big[\int_t^{t+\delta t} e^{-r(s-t)}\big[c(s)-a(c(s))\big]P(s)ds + \int_{t+\delta t}^{T} e^{-r(s-t)}\big[c(s)-a(c(s))\big]P(s)ds \\
&\qquad\qquad + e^{-r(T-t)}\hat{V}\big(P(T),I(T),T\big)\Big] \\
&= \sup_{c(s)\in C(I(s))} E^{\mathbb{Q}}\Big[\int_t^{t+\delta t} e^{-r(s-t)}\big[c(s)-a(c(s))\big]P(s)ds \\
&\qquad\qquad + e^{-r\delta t}\Big[\int_{t+\delta t}^{T} e^{-r(s-(t+\delta t))}\big[c(s)-a(c(s))\big]P(s)ds + e^{-r(T-(t+\delta t))}\hat{V}\big(P(T),I(T),T\big)\Big]\Big] \\
&= \sup_{c(s)\in C(I(s))} E^{\mathbb{Q}}\left[\int_t^{t+\delta t} e^{-r(s-t)}\big[c(s)-a(c(s))\big]P(s)ds + e^{-r\delta t}\hat{V}\big(P+\delta P, I+\delta I, t+\delta t\big)\right],
\end{aligned}
\tag{A.1}
$$

where we use dynamic programming to obtain the last equality and where δP, δI are the random changes from (P,I) in the interval $t \to t+\delta t$. Using notation $\delta\hat{V} = \hat{V}(P+\delta P, I+$

$\delta I, t+\delta t) - \hat{V}(P, I, t)$ we can write the above equation as

$$\sup_{c(s)\in C(I(s))} E^{\mathbb{Q}}\left[\int_t^{t+\delta t} e^{-r(s-t)}\big[c(s)-a(c(s))\big]P(s)ds + e^{-r\delta t}\delta\hat{V} + \big(e^{-r\delta t}-1\big)\hat{V}\right] = 0. \tag{A.2}$$

Using Ito's Lemma as well as equations (2.2.4-2.2.5) and (2.2.9) give

$$\begin{aligned}\delta\hat{V} = &\int_t^{t+\delta t} U(s)ds + \int_t^{t+\delta t} \hat{\sigma}(P(s))P(s)\hat{V}_P(P(s), I(s), s)dZ(s) \\ &- \int_t^{t+\delta t}\big[c(s)+a(c(s))\big]\hat{V}_I(P(s), I(s), s)ds,\end{aligned} \tag{A.3}$$

where $U(s)$ is defined as

$$\begin{aligned}U(s) = &\hat{V}_t(P(s), I(s), s) + \alpha\big[K(s)-P(s)\big]\hat{V}_P(P(s), I(s), s) \\ &+ \frac{1}{2}\hat{\sigma}^2(P(s))P^2(s)\hat{V}_{PP}(P(s), I(s), s)\end{aligned} \tag{A.4}$$

Substituting (A.3) into (A.2) and using $E^{\mathbb{Q}}[dZ] = 0$ yields

$$\begin{aligned}\sup_{c(s)\in C(I(s))} E^{\mathbb{Q}}\Bigg[&\int_t^{t+\delta t} e^{-r(s-t)}\big[c(s)-a(c(s))\big]P(s)ds \\ &+ e^{-r\delta t}\Big[\int_t^{t+\delta t} U(s)ds - \int_t^{t+\delta t}\big[c(s)+a(c(s))\big]\hat{V}_I ds\Big] + \big(e^{-r\delta t}-1\big)\hat{V}\Bigg] = 0.\end{aligned} \tag{A.5}$$

If we take a constant control $c(s) = \bar{c}$ for $t \le s \le t+\delta t$, where $\bar{c}$ is any value in $C(I(t))$, then (A.5) implies

$$\begin{aligned}E^{\mathbb{Q}}\Bigg[&\int_t^{t+\delta t} e^{-r(s-t)}(\bar{c}-a(\bar{c}))P(s)ds + e^{-r\delta t}\Big[\int_t^{t+\delta t} U(s)ds - \int_t^{t+\delta t}(\bar{c}+a(\bar{c}))\hat{V}_I ds\Big] \\ &+ \big(e^{-r\delta t}-1\big)\hat{V}\Bigg] \le 0.\end{aligned} \tag{A.6}$$

Dividing through by δt and taking limits as $\delta t \to 0$ (assuming we can interchange the

limit with the expectation) leads to

$$\hat{V}_t + \alpha(K(t) - P)\hat{V}_P + \frac{1}{2}\hat{\sigma}^2(P)P^2\hat{V}_{PP} + (\bar{c} - a(\bar{c}))P - (\bar{c} + a(\bar{c}))\hat{V}_I - r\hat{V} \leq 0. \quad \text{(A.7)}$$

On the other hand, we assume that $c^*(s)$ is a continuous optimal control path that achieves the supremum in (A.5). Then following a similar argument gives

$$\hat{V}_t + \alpha(K(t) - P)\hat{V}_P + \frac{1}{2}\hat{\sigma}^2(P)P^2\hat{V}_{PP} + \big[c^*(t) - a(c^*(t))\big]P - \big[c^*(t) + a(c^*(t))\big]\hat{V}_I - r\hat{V} = 0. \quad \text{(A.8)}$$

Consequently, inequality (A.7) holds for all controls $c \in C(I)$, while it holds with equality for $c = c^*(t)$ (according to (A.8)), which is equivalent to the following HJB equation

$$\sup_{c \in C(I)} \Big[\hat{V}_t + \alpha(K(t) - P)\hat{V}_P + \frac{1}{2}\hat{\sigma}^2(P)P^2\hat{V}_{PP} + (c - a(c))P - (c + a(c))\hat{V}_I - r\hat{V}\Big] = 0. \quad \text{(A.9)}$$

Moving all terms independent of the control c out of the sup expressions result in

$$\hat{V}_t + \alpha(K(t) - P)\hat{V}_P + \frac{1}{2}\hat{\sigma}^2(P)P^2\hat{V}_{PP} + \sup_{c \in C(I)} \big[(c - a(c))P - (c + a(c))\hat{V}_I\big] - r\hat{V} = 0. \quad \text{(A.10)}$$

Appendix B

Discrete Equation Coefficients

In this appendix, we give the coefficients of discrete differential operator $\mathcal{L}_h$ defined in (2.3.4). For $i = 0$, we impose boundary condition (2.2.16) as $P \to 0$ by using forward differencing to evaluate the first-order derivative term in (2.3.2) and setting $V_{PP} = 0$, which results in

$$\gamma_0^n = 0 \quad ; \quad \beta_0^n = \frac{\alpha(K(T-\tau^n) - P_0)}{P_1 - P_0}. \tag{B.1}$$

Similarly, for $i = i_{\max}$, the boundary condition (2.2.17) at $P = P_{\max}$ can be imposed by setting

$$\gamma_{i_{\max}}^n = -\frac{\alpha(K(T-\tau^n) - P_{i_{\max}})}{P_{i_{\max}} - P_{i_{\max}-1}} \quad ; \quad \beta_{i_{\max}}^n = 0, \tag{B.2}$$

where $\gamma_{i_{\max}}^n$ and $\beta_{i_{\max}}^n$ are obtained by evaluating the first-order derivative term in (2.3.2) with backward differencing and setting $V_{PP} = 0$.

Away from $i = 0$ and $i = i_{\max}$, applying the second-order central differencing in the first and second order derivative terms in equation (2.3.2) leads to the following values of coefficients γ_i^n and β_i^n:

$$\begin{aligned} \gamma_{i,central}^n &= \frac{(\hat{\sigma}(P_i)P_i)^2}{(P_i - P_{i-1})(P_{i+1} - P_{i-1})} - \frac{\alpha(K(T-\tau^n) - P_i)}{P_{i+1} - P_{i-1}}, \\ \beta_{i,central}^n &= \frac{(\hat{\sigma}(P_i)P_i)^2}{(P_{i+1} - P_i)(P_{i+1} - P_{i-1})} + \frac{\alpha(K(T-\tau^n) - P_i)}{P_{i+1} - P_{i-1}}, \quad i = 1, \ldots, i_{\max} - 1. \end{aligned} \tag{B.3}$$

If either $\gamma_{i,central}^n$ or $\beta_{i,central}^n$ is negative, the discrete scheme will not be monotone. Mono-

tonicity can be restored by using first-order forward or backward differencing in the first-order derivative term at the problem nodes. Forward differencing produces:

$$\begin{aligned} \gamma^n_{i,forward} &= \frac{(\hat{\sigma}(P_i)P_i)^2}{(P_i - P_{i-1})(P_{i+1} - P_{i-1})}, \\ \beta^n_{i,forward} &= \frac{(\hat{\sigma}(P_i)P_i)^2}{(P_{i+1} - P_i)(P_{i+1} - P_{i-1})} + \frac{\alpha(K(T-\tau^n) - P_i)}{P_{i+1} - P_i}, \quad i = 1, \ldots, i_{\max} - 1. \end{aligned} \tag{B.4}$$

while backward differencing gives:

$$\begin{aligned} \gamma^n_{i,backward} &= \frac{(\hat{\sigma}(P_i)P_i)^2}{(P_i - P_{i-1})(P_{i+1} - P_{i-1})} - \frac{\alpha(K(T-\tau^n) - P_i)}{P_i - P_{i-1}}, \\ \beta^n_{i,backward} &= \frac{(\hat{\sigma}(P_i)P_i)^2}{(P_{i+1} - P_i)(P_{i+1} - P_{i-1})}, \quad i = 1, \ldots, i_{\max} - 1. \end{aligned} \tag{B.5}$$

On one hand, we want to use central differencing to achieve second-order correctness. On the other hand, we need to maintain γ^n_i and β^n_i positive so that the scheme is monotone. Consequently, we decide on central or forward/backward discretization at each node (P_i, I_j), $i = 1, \ldots, i_{\max} - 1$, $j = 0, \ldots, j_{\max}$, based on Algorithm B.1.

If $[\gamma^n_{i,central} \geq 0$ and $\beta^n_{i,central} \geq 0]$ then
  $\gamma^n_i = \gamma^n_{i,central}$
  $\beta^n_i = \beta^n_{i,central}$
ElseIf $[\beta^n_{i,forward} \geq 0]$ then
  $\gamma^n_i = \gamma^n_{i,forward}$
  $\beta^n_i = \beta^n_{i,forward}$
Else
  $\gamma^n_i = \gamma^n_{i,backward}$
  $\beta^n_i = \beta^n_{i,backward}$
EndIf

ALGORITHM B.1: *Selection of the central, forward or backward discretization*

Note that (B.1-B.2) and Algorithm B.1 guarantee that both γ^n_i and β^n_i are non-negative. Equations (B.3) imply that $\gamma^n_{i,central}$ and $\beta^n_{i,central}$ are positive if P_i is close to $K(T - \tau^n)$, which is the equilibrium price of the risk neutral natural gas spot price at

$t = T - \tau^n$. Hence the strategy in Algorithm B.1 will use central differencing for those nodes P_i close to the equilibrium price. These nodes are in the region of interest, since the gas spot price will not stray too far away from the mean-reversion equilibrium price. Consequently, the use of a low-order differencing scheme for nodes far away from the equilibrium price should not result in poor convergence for the nodes near the equilibrium price.

Appendix C

Discrete Optimal Control Strategy for the Natural Gas Storage

In this Appendix, we derive a discretization based on purely physical reasoning, assuming that the operator of a gas storage facility can change the controls only at fixed discrete times. We will see that the final discrete equations are identical to the fully implicit semi-Lagrangian discretization derived in Section 2.3.2. This provides justification for the choice of integration points used in deriving the fully implicit discretization (2.3.15-2.3.17). It is interesting to observe that the fully implicit semi-Lagrangian discretization can be interpreted as a no-arbitrage jump condition.

Suppose we have $N+1$ discrete times, denoted by $0 = t^0 < t^1 < \ldots < t^N = T$, over the period from $t = 0$ to $t = T$. Let $t^0, t^1, \ldots, t^{N-1}$ be the discrete operation times where the operator of a natural gas storage facility can choose to change the controls. At each decision time $t = t^{n-1}$, $n = 1, 2, \ldots, N$, the operator either does nothing or chooses to inject or produce gas at a constant rate. This decision cannot be changed until the next decision time t^n.

We assume that no cash flows will occur until the end of a time interval. In other words, given a decision at t^{n-1}, then the revenue is obtained for all the gas produced in $[t^{n-1}, t^n)$ at t^n. We also imagine that the inventory is regarded as constant during $[t^{n-1}, t^n)$. The inventory is adjusted only at t^n to reflect the gas produced/injected during

$[t^{n-1}, t^n)$. Clearly, if we let $N \to \infty$, then the above discrete control scenario will turn into a continuous control scenario.

Let us consider a path of inventory level in the gas storage facility at different times with respect to a fixed natural gas spot price, denoted by $\hat{\mathcal{I}}(t)$. Such a path is generated by the control strategy that the operator applies to the storage facility along a specific time path.[1] As described above, under our discrete control scenario, the gas inventory changes only at discrete times t^n. Thus, the path $\hat{\mathcal{I}}(t)$ is a piece-wise constant function with $\hat{\mathcal{I}}(t) = \hat{\mathcal{I}}(t^{n-1})$ for $t \in [t^{n-1}, t^n)$, $n = 1, 2, \ldots, N$. Let $\hat{V}\big(P, \hat{\mathcal{I}}(t), t\big)$ denote the value of the gas storage facility at forward time t with gas spot price fixed at P and gas inventory following the path $\hat{\mathcal{I}}(t)$. In the rest of this section, we will investigate the value of $\hat{V}\big(P, \hat{\mathcal{I}}(t), t\big)$ at discrete times $t = t^{n-1}$, $n = 1, 2, \ldots, N$.[2] Note that in contrast to the inventory path $\hat{\mathcal{I}}$, we use notation I to represent a gas inventory value independent of P and t, as appears in PDE (2.2.12).

We denote by $\hat{\zeta}\big(P, \hat{\mathcal{I}}(t^{n-1}), t^{n-1}\big)$ the constant rate of the control chosen by the storage facility operator at $t = t^{n-1}$ when the gas spot price is fixed at P and gas inventory path $\hat{\mathcal{I}}$ arrives at $\hat{\mathcal{I}}(t^{n-1})$. The control $\hat{\zeta}\big(P, \hat{\mathcal{I}}(t^{n-1}), t^{n-1}\big)$ is applied to the gas storage facility for $t \in [t^{n-1}, t^n)$. In accordance with the convention in Section 2.2.1, we assume that $\hat{\zeta}\big(P, \hat{\mathcal{I}}(t^{n-1}), t^{n-1}\big) > 0$ represents production, $\hat{\zeta}\big(P, \hat{\mathcal{I}}(t^{n-1}), t^{n-1}\big) < 0$ represents injection, and $\hat{\zeta}\big(P, \hat{\mathcal{I}}(t^{n-1}), t^{n-1}\big) = 0$ represents no operation. From Section 2.2.1, we require the control $\hat{\zeta}$ to satisfy the constraint $\hat{\zeta}\big(P, \hat{\mathcal{I}}(t^{n-1}), t^{n-1}\big) \in C(\hat{\mathcal{I}}(t^{n-1}))$, where $C(\hat{\mathcal{I}}) = [c_{\min}(\hat{\mathcal{I}}), -k_5] \cup [0, c_{\max}(\hat{\mathcal{I}})]$, as defined in equation (2.2.7).

According to the above assumptions, the gas inventory along the path $\hat{\mathcal{I}}$ is fixed at $\hat{\mathcal{I}}(t^{n-1})$ for $t \in [t^{n-1}, t^n)$, and switches to $\hat{\mathcal{I}}(t^n)$ at $t = t^n$. Let $\Delta t = t^n - t^{n-1}$. The change of the gas inventory from $t = t^{n-1}$ to $t = t^n$, due to gas injection, production and gas loss,

[1] We do not need to introduce the control as a variable in $\hat{\mathcal{I}}$ because we assume that an optimal control strategy will always be chosen by a rational operator, as explained below, under our discrete control scenario; the optimal strategy will produce an optimal control that is determined by the gas spot price P and gas inventory $\hat{\mathcal{I}}$ at the decision time. Hence, we will not set the control as a variable in $\hat{\mathcal{I}}$.

[2] Note that $\hat{V}\big(P, \hat{\mathcal{I}}(t), t\big)$ for $t = t^N$ is given by a payoff condition.

satisfies

$$\hat{\mathcal{I}}(t^n) = \hat{\mathcal{I}}(t^{n-1}) - \Delta t \left[\hat{\zeta}\big(P, \hat{\mathcal{I}}(t^{n-1}), t^{n-1}\big) + a\left(\hat{\zeta}\big(P, \hat{\mathcal{I}}(t^{n-1}), t^{n-1}\big)\right)\right]. \tag{C.1}$$

Note that equation (C.1) is the discrete-time version of the ODE (2.2.4) with variable c replaced by $\hat{\zeta}$. For convenience, we introduce the notation $\hat{\mathcal{I}}^n = \hat{\mathcal{I}}(t^n)$, $\hat{\zeta}^n = \hat{\zeta}\big(P, \hat{\mathcal{I}}(t^n), t^n\big)$. We drop the dependence of $\hat{\zeta}$ on $(P, \hat{\mathcal{I}})$ for the purpose of simplifying our presentation. Using the notation introduced above, equation (C.1) can be written as

$$\hat{\mathcal{I}}^n = \hat{\mathcal{I}}^{n-1} - \Delta t\big(\hat{\zeta}^{n-1} + a(\hat{\zeta}^{n-1})\big). \tag{C.2}$$

Due to physical storage constraints, after imposing the control $\hat{\zeta}^{n-1}$ at $t = t^{n-1}$, the resulting gas inventory $\hat{\mathcal{I}}^n$ from (C.2) must remain inside the domain $[0, I_{\max}]$. In particular, if $\hat{\mathcal{I}}^{n-1} = 0$, then the operator cannot produce gas at $t = t^{n-1}$ and thus we require $\hat{\zeta}^{n-1} + a(\hat{\zeta}^{n-1}) \leq 0$. Similarly, if $\hat{\mathcal{I}}^{n-1} = I_{\max}$, we require $\hat{\zeta}^{n-1} + a(\hat{\zeta}^{n-1}) \geq 0$. These two conditions are the discrete version of boundary conditions (2.2.14) and (2.2.15) satisfied by PDE (2.2.12), respectively. We refer to a control $\hat{\zeta}^{n-1}$ as admissible control if it satisfies $\hat{\zeta}^{n-1} \in C(\hat{\mathcal{I}}^{n-1})$ and $\hat{\mathcal{I}}^n$ calculated from (C.2) is bounded within $[0, I_{\max}]$. We denote by $\hat{C}^{n-1}$ the set of all admissible controls at $t = t^{n-1}$.

Recall that, under the discrete control assumptions, for $t \in [t^{n-1}, t^n)$, the gas inventory is fixed at $\hat{\mathcal{I}}^{n-1}$ and no cash flows appear. At $t = t^n$, the inventory switches to $\hat{\mathcal{I}}^n$ and a revenue $\Delta t(\hat{\zeta}^{n-1} - a(\hat{\zeta}^{n-1}))P$ is created, which is the product of the amount of gas lost, injected or produced, i.e., $\Delta t(\hat{\zeta}^{n-1} - a(\hat{\zeta}^{n-1}))$, during the period of $[t^{n-1}, t^n)$ and the gas spot price. Let $t^{n-} = t^n - \epsilon$ with $\epsilon > 0$, $\epsilon \to 0$, that is, t^{n-} is the time infinitesimally before t^n. By no-arbitrage [86], at $t = t^n$, $\hat{V}$ must satisfy the following condition

$$\hat{V}(P, \hat{\mathcal{I}}^n, t^n) = \hat{V}(P, \hat{\mathcal{I}}^{n-1}, t^{n-}) - \Delta t\big(\hat{\zeta}^{n-1} - a(\hat{\zeta}^{n-1})\big)P. \tag{C.3}$$

Out of a variety of choices, the problem next is to find the optimal value of $\hat{\zeta}^{n-1}$. A rational operator will choose a control $\hat{\zeta}^{n-1}$ that maximizes the facility value $\hat{V}$ at $t = t^{n-}$. Thus,

the no-arbitrage condition (C.3) can be written as

$$\hat{V}(P,\hat{\mathcal{I}}^{n-1},t^{n-}) = \sup_{\hat{\zeta}^{n-1}\in\hat{C}^{n-1}} \left\{ \hat{V}(P,\hat{\mathcal{I}}^{n},t^{n}) + \Delta t\big(\hat{\zeta}^{n-1} - a(\hat{\zeta}^{n-1})\big)P \right\}, \tag{C.4}$$

$$\hat{\mathcal{I}}^{n} = \hat{\mathcal{I}}^{n-1} - \Delta t\big(\hat{\zeta}^{n-1} + a(\hat{\zeta}^{n-1})\big) \;\; ; \;\; \hat{\mathcal{I}}^{n} \in [0, I_{\max}] \text{ for any } \hat{\zeta}^{n-1} \in \hat{C}^{n-1}. \tag{C.5}$$

During the interval $[t^{n-1}, t^{n-}]$, the inventory in the path $\hat{\mathcal{I}}$ is a constant $\hat{\mathcal{I}}^{n-1}$ and no cash flows occur, hence the value of the facility is given by the solution of the PDE

$$\hat{V}_t + \frac{1}{2}\hat{\sigma}^2(P)P^2\hat{V}_{PP} + \alpha(K(t)-P)\hat{V}_P - r\hat{V} = 0 \; . \tag{C.6}$$

We can semidiscretize this in time over the interval $[t^{n-1}, t^{n-}]$ using fully implicit timestepping

$$\begin{aligned} &\frac{\hat{V}(P,\hat{\mathcal{I}}^{n-1},t^{n-}) - \hat{V}(P,\hat{\mathcal{I}}^{n-1},t^{n-1})}{\Delta t} + \frac{\hat{\sigma}^2(P)P^2}{2}\hat{V}_{PP}(P,\hat{\mathcal{I}}^{n-1},t^{n-1}) \\ &\quad + \alpha\big(K(t^{n-1}) - P\big)\hat{V}_P(P,\hat{\mathcal{I}}^{n-1},t^{n-1}) - r\hat{V}(P,\hat{\mathcal{I}}^{n-1},t^{n-1}) = 0 \;\; . \end{aligned} \tag{C.7}$$

Note that we will solve equation (C.6) backward in time, so that fully implicit timestepping uses the values of the diffusion, drift and discounting terms at $t = t^{n-1}$.

Substituting equation (C.4) into equation (C.7) gives

$$\begin{aligned} &\frac{\sup_{\hat{\zeta}^{n-1}\in\hat{C}^{n-1}} \left\{ \hat{V}(P,\hat{\mathcal{I}}^{n},t^{n}) + \Delta t\big(\hat{\zeta}^{n-1} - a(\hat{\zeta}^{n-1})\big)P \right\} - \hat{V}(P,\hat{\mathcal{I}}^{n-1},t^{n-1})}{\Delta t} \\ &+ \frac{\hat{\sigma}^2(P)P^2}{2}\hat{V}_{PP}(P,\hat{\mathcal{I}}^{n-1},t^{n-1}) + \alpha\big(K(t^{n-1}) - P\big)\hat{V}_P(P,\hat{\mathcal{I}}^{n-1},t^{n-1}) \\ &- r\hat{V}(P,\hat{\mathcal{I}}^{n-1},t^{n-1}) = 0 \;\; . \end{aligned} \tag{C.8}$$

Based on the semidiscretization scheme (C.8), assuming $\hat{V}(P,\hat{\mathcal{I}}^{n},t^{n})$ is known, we are able to compute $\hat{V}(P,\hat{\mathcal{I}}^{n-1},t^{n-1})$ at the previous time $t = t^{n-1}$ with gas inventory moving along a path (given in (C.5)) generated by the optimal control strategy.

Equation (C.8) is written in terms of forward times t. We can rewrite this equation in terms of backward times $\tau = T - t$. Let $\mathcal{I}(\tau)$ denote the gas inventory path at backward

times as a function of τ, that is, $\mathcal{I}(\tau) = \mathcal{I}(T-t) = \hat{\mathcal{I}}(t)$. Let $\zeta\big(P,\mathcal{I}(\tau),\tau\big)$ denote the control at backward times as a function of $P,\mathcal{I},\tau$ with $\zeta\big(P,\mathcal{I}(\tau),\tau\big) = \zeta\big(P,\mathcal{I}(T-t),T-t\big) = \hat{\zeta}\big(P,\hat{\mathcal{I}}(t),t\big)$. Similarly, we can define the value of a gas storage facility as a function of $P,\mathcal{I},\tau$ as $V\big(P,\mathcal{I}(\tau),\tau\big)$, where, in terms of $\hat{V}\big(P,\hat{\mathcal{I}},t\big)$, we have the identity $V\big(P,\mathcal{I}(\tau),\tau\big) = V\big(P,\mathcal{I}(T-t),T-t\big) = \hat{V}\big(P,\hat{\mathcal{I}}(t),t\big)$. Let $k = N-n$, $k = 0,1,\ldots,N-1$, so that k counts backwards. Since $t^N = T$, we have that $\tau^k = T - t^n$, $\tau^{k+1} = T - t^{n-1}$. Let $\mathcal{I}^k = \mathcal{I}(\tau^k)$, $\zeta^k = \zeta\big(P,\mathcal{I}(\tau^k),\tau^k\big)$, $\Delta\tau = \tau^{k+1} - \tau^k$.[3] We can obtain the following identities between discrete forward and backward times:

$$\hat{\zeta}^{n-1} = \zeta^{k+1}; \quad \hat{\mathcal{I}}^{n-1} = \mathcal{I}^{k+1}; \quad \hat{\mathcal{I}}^{n} = \mathcal{I}^{k}; \quad \Delta t = \Delta\tau; \quad K(t^{n-1}) = K(T-\tau^{k+1}) \tag{C.9}$$
$$\hat{V}(P,\hat{\mathcal{I}}^{n-1},t^{n-1}) = V(P,\mathcal{I}^{k+1},\tau^{k+1}); \quad \hat{V}(P,\hat{\mathcal{I}}^{n},t^{n}) = V(P,\mathcal{I}^{k},\tau^{k}).$$

Using identities of (C.9) in equation (C.8), we obtain the optimal control problem in backwards time

$$\begin{aligned}
&\frac{V(P,\mathcal{I}^{k+1},\tau^{k+1}) - \sup_{\zeta^{k+1}\in C^{k+1}}\left\{V(P,\mathcal{I}^{k},\tau^{k}) + \Delta\tau(\zeta^{k+1} - a(\zeta^{k+1}))P\right\}}{\Delta\tau} \\
&= \frac{\hat{\sigma}^2(P)P^2}{2}V_{PP}(P,\mathcal{I}^{k+1},\tau^{k+1}) + \alpha\big(K(T-\tau^{k+1}) - P\big)V_P(P,\mathcal{I}^{k+1},\tau^{k+1}) \\
&\quad - rV(P,\mathcal{I}^{k+1},\tau^{k+1}),
\end{aligned} \tag{C.10}$$

where

$$\mathcal{I}^k = \mathcal{I}^{k+1} - \Delta\tau\big(\zeta^{k+1} + a(\zeta^{k+1})\big) \;\; ; \;\; C^{k+1} \subseteq C(\mathcal{I}^{k+1}) \;,\; \mathcal{I}^k \in [0, I_{\max}] \text{ for any } \zeta^{k+1} \in C^{k+1} \;. \tag{C.11}$$

Let us further discretize the scheme (C.11-C.10) in the $(P,\mathcal{I})$ direction by setting

$$k = n, \quad P = P_i, \quad \mathcal{I}^{k+1} = I_j, \quad \mathcal{I}^k = I^n_{\hat{j}}, \quad \zeta^{k+1} = \zeta^{n+1}_{i,j}, \quad C^{k+1} = C^{n+1}_j,$$
$$V(P,\mathcal{I}^{k+1},\tau^{k+1}) = V^{n+1}_{i,j}, \quad V(P,\mathcal{I}^{k},\tau^{k}) = V^{n}_{i,\hat{j}}$$

in (C.11-C.10) and replacing the right hand side of (C.10) by the discrete differential

[3] Here we also drop the dependence of ζ on $(P,\mathcal{I})$ for convenience.

operator $(\mathcal{L}_h V)_{i,j}^{n+1}$ given in (2.3.4). Then we can recognize that the full discretization of the semi-discretization (C.11-C.10) (based on a discrete optimal control) is algebraically identical to the semi-Lagrangian, fully implicit discretization (2.3.15-2.3.17).

Remark C.1. *The optimal control in (C.10) is evaluated at $\tau = \tau^{k+1}$ (i.e., at the end of the timestep, going backwards). Note that this corresponds to the beginning of a time interval going forwards in time. This is because, as is usual in stochastic control, the operator must make a decision at the beginning of a time interval, and then let the system evolve randomly.*

Appendix D

Consistency Proof for the Gas Storage Problem

Prior to proving Lemma 3.19, we first show the following Lemmas.

Lemma D.1. *Let $C_j^{n+1} \subseteq C\big(I_{\hat{j}}^n\big) \times C(I_j)$ be the set of admissible controls such that $I_{\hat{j}}^n$ calculated from*

$$I_{\hat{j}}^n = I_j - (1-\theta)\Delta\tau\big(\zeta_{i,j}^{n+1} + a(\zeta_{i,j}^{n+1})\big) - \theta\Delta\tau\big(\zeta_{i,j}^{n} + a(\zeta_{i,j}^{n})\big) \tag{D.1}$$

is bounded within $[0, I_{\max}]$, where $C\big(I_{\hat{j}}^n\big), C(I_j)$ are defined based on equation (2.2.7). Then by taking $\Delta\tau$ sufficiently small, we have

$$C_j^{n+1} = C\big(I_{\hat{j}}^n\big) \times C(I_j) \tag{D.2}$$

provided that $c_{\max}$, $c_{\min}$ and $a(c)$ satisfy equations (2.2.1), (2.2.3) and (2.2.5), respectively.

Proof. Consider the case where $\epsilon^* \le I_j \le I_{\max} - \epsilon^*$, where $\epsilon^* \ll I_{\max}$ is independent of h. In other words, we exclude a small strip of finite size near $I = 0$ and $I = I_{\max}$.

If we let $\zeta_{i,j}^n, \zeta_{i,j}^{n+1}$ be two arbitrary controls satisfying

$$\zeta_{i,j}^n \in C\big(I_{\hat{j}}^n\big) \quad ; \quad \zeta_{i,j}^{n+1} \in C(I_j), \tag{D.3}$$

then since $\zeta_{i,j}^n, \zeta_{i,j}^{n+1}$ are bounded, there exists a constant $D > 0$ such that

$$|\zeta_{i,j}^n + a\big(\zeta_{i,j}^n\big)| < D \quad ; \quad |\zeta_{i,j}^{n+1} + a\big(\zeta_{i,j}^{n+1}\big)| < D. \tag{D.4}$$

Equations (D.1) and (D.4) imply that

$$I_j - \Delta\tau D < I_{\hat{j}}^n < I_j + \Delta\tau D. \tag{D.5}$$

Thus, by taking $\Delta\tau < \min\big\{\frac{I_j}{D}, \frac{I_{\max}-I_j}{D}\big\}$, from equation (D.5), we obtain that $I_{\hat{j}}^n \in [0, I_{\max}]$, assuming $\epsilon^* \leq I_j \leq I_{\max} - \epsilon^*$. The above argument shows that $\zeta_{i,j}^n$ and $\zeta_{i,j}^{n+1}$ are admissible controls such that $(\zeta_{i,j}^n, \zeta_{i,j}^{n+1}) \in C_j^{n+1}$ for a sufficiently small $\Delta\tau$. This amounts to the identities in (D.2) for the case $\epsilon^* \leq I_j \leq I_{\max} - \epsilon^*$.

Now, consider the nodes in the strip $I_j < \epsilon^*$. For j fixed, we have $I_j = O(h)$, hence the condition on $\Delta\tau$ becomes

$$\Delta\tau \quad \leq \quad const.\frac{h}{D} \quad , \tag{D.6}$$

which implies that we have consistency only if there is a condition on $\Delta\tau/h$. Recall that we have assumed only that condition (2.3.1) holds for any constant C_3. Condition (D.6) adds an extra condition which seems unnatural. The strip $I_j > I_{\max} - \epsilon^*$ also gives rise to a similar condition.

In fact, this condition disappears if we carry out a more precise analysis. This proof is very long and tedious, and involves examination of many different cases. We will outline the proof for a simple case only.

Consider the fully implicit case, and $I_j < \epsilon^*$. In this case we have

$$I_{\hat{j}}^n \quad = \quad I_j - \Delta\tau\big(\zeta_{i,j}^{n+1} + a\big(\zeta_{i,j}^{n+1}\big)\big). \tag{D.7}$$

Equations (2.2.1), (2.2.5) and (2.2.6) imply that

$$I_j - \Delta\tau c_{\max}(I_j) \quad \leq \quad I_{\hat{j}}^n \tag{D.8}$$

so that in order to have $I_{\hat{j}}^n > 0$, we must have

$$\Delta\tau c_{\max}(I_j) \;\; < \;\; I_j \tag{D.9}$$

From equation (2.2.1), we have that

$$c_{\max}(I) = O(\sqrt{I}) \;\; ; \;\; I \to 0 \;\; . \tag{D.10}$$

Assuming that $I_j = O(h)$, then equations (D.9-D.10) give

$$\Delta\tau \;\; = \;\; O(\sqrt{h}) \; , \tag{D.11}$$

but this is a weaker condition than the assumption (2.3.1). Hence condition (D.9) is satisfied for any $\zeta_{i,j}^{n+1} \in C(I_j)$ when $\Delta\tau$ sufficiently small, with no additional restriction on $\Delta\tau/h$ (other than assumption (2.3.1)). Equation (D.2) holds in this case. Similar arguments show that (D.2) holds in all cases for $0 \leq I_j \leq I_{\max}$, for $\Delta\tau$ sufficiently small, provided that $c_{\max}$, $c_{\min}$ and $a(c)$ satisfy equations (2.2.1), (2.2.3) and (2.2.5), respectively. □

Lemma D.2. *Suppose $F(c_1, c_2)$ and $H(c_1, c_2)$ are bounded functions defined in some domain $c_1 \in C_1$, $c_2 \in C_2$. Then there exists a bounded function $Q(h)$ where*

$$Q(h) \leq \sup_{c_1,c_2} |H(c_1, c_2)| \tag{D.12}$$

such that

$$\inf_{c_1,c_2} \{F(c_1, c_2) + H(c_1, c_2)h\} = \inf_{c_1,c_2} \{F(c_1, c_2)\} + hQ(h), \tag{D.13}$$

Proof. According to [46], we have that

$$\inf_{c_1,c_2} \{H(c_1, c_2)h\} \leq \inf_{c_1,c_2} \{F(c_1, c_2) + H(c_1, c_2)h\} - \inf_{c_1,c_2} \{F(c_1, c_2)\} \leq \sup_{c_1,c_2} \{H(c_1, c_2)h)\} \, . \tag{D.14}$$

Equation (D.13) directly follows from the above inequalities. □

Lemma D.3 (Point-wise Consistency)**.** *Assume that all conditions in Lemma 3.19 are satisfied. Let* $\mathbf{x} = (P_i, I_j, \tau^{n+1})$*. Then for any test function* $\phi(P, I, \tau)$ *having bounded derivatives of all orders in* $(P, I, \tau) \in \bar{\Omega}$ *with* $\phi_{i,j}^{n+1} = \phi(P_i, I_j, \tau^{n+1})$*, we have*

$$\begin{aligned} \lim_{h\to 0}\Big|\bar{F}\big(D^2\phi(\mathbf{x}), D\phi(\mathbf{x}), \phi(\mathbf{x}), \mathbf{x}\big) - \mathcal{G}_{i,j}^{n+1}\big(h, \phi_{i,j}^{n+1}, \{\phi_{k,j}^{n+1}\}_{k\neq i}, \{\phi_{i,j}^{n}\}\big)\Big| = 0; \\ i = 0, \ldots, i_{\max}, \quad j = 0, \ldots, j_{\max}, \quad n+1 = 0, \ldots, N\ , \end{aligned} \tag{D.15}$$

where the function $\bar{F}$ *is given in (3.1.39).*

Proof. When $n+1 = 0$, $\bar{F} = \phi(\mathbf{x}) - B(\mathbf{x})$. Then according to equation (3.2.2), it is obvious to see that equation (D.15) holds with $n+1 = 0$.

We then consider the case when $n+1 \geq 1$. In this case,

$$\begin{aligned} &\bar{F}\big(D^2\phi(\mathbf{x}), D\phi(\mathbf{x}), \phi(\mathbf{x}), \mathbf{x}\big) \\ &= \inf_{c\in C(I_j)} \Big\{\big(\phi_\tau + (c + a(c))\phi_I - (c - a(c))P - \mathcal{L}\phi\big)_{i,j}^{n+1}\Big\}, \quad \text{if } n+1 \geq 1. \end{aligned} \tag{D.16}$$

Let us consider the discrete function $\mathcal{G}_{i,j}^{n+1}$ in equation (3.2.1) for $n+1 \geq 1$. We define vector ϕ^n with components $[\phi^n]_{i,j} = \phi(P_i, I_j, \tau^n)$. Assuming the discretization in Appendix B is used, then from Taylor series expansions, we obtain that

$$(L\phi^{n+1})_{i,j} = (\mathcal{L}\phi)_{i,j}^{n+1} + O(\Delta P_{\max}). \tag{D.17}$$

Let us assume that the interpolation error in equation (2.3.23), due to operation Φ^{n+1}, is $O((\Delta I_{\max})^2)$. In other words, linear or higher order interpolation is used to compute $\phi\big(P_i, I_{\hat{j}}^n, \tau^n\big)$. Consequently, we can write

$$\big[\Phi^{n+1}\phi^n\big]_{i,j} = \phi\big(P_i, I_{\hat{j}}^n, \tau^n\big) + O((\Delta I_{\max})^2) \tag{D.18}$$

$$\big[\Phi^{n+1}L\phi^n\big]_{i,j} = (\mathcal{L}\phi)\big(P_i, I_{\hat{j}}^n, \tau^n\big) + O(\Delta P_{\max} + (\Delta I_{\max})^2), \tag{D.19}$$

where

$$I_{\hat{j}}^n = I_j - (1-\theta)\Delta\tau\big(\zeta_{i,j}^{n+1} + a(\zeta_{i,j}^{n+1})\big) - \theta\Delta\tau\big(\zeta_{i,j}^{n} + a(\zeta_{i,j}^{n})\big), \tag{D.20}$$

and $\mathcal{L}\phi$ is regarded as a single function. In the above equation (D.19), we use (D.17) to estimate the error between the discrete operation $[L\phi^n]_{ij}$, as given in (2.3.22), and the exact operation $(\mathcal{L}\phi)^n_{i,j}$.

According to (D.20) and (D.18), we have that

$$\begin{aligned}\frac{\phi^{n+1}_{i,j} - \left[\Phi^{n+1}\phi^n\right]_{i,j}}{\Delta\tau} &= (\phi_\tau)^{n+1}_{i,j} + (1-\theta)\big(\zeta^{n+1}_{i,j} + a(\zeta^{n+1}_{i,j})\big)(\phi_I)^n_{i,j} + \theta\big(\zeta^n_{i,j} + a(\zeta^n_{i,j})\big)(\phi_I)^n_{i,j} \\ &\quad + O\left(\Delta\tau + (\Delta I_{\max})^2/\Delta\tau\right),\end{aligned} \tag{D.21}$$

where Taylor series is used to expand $\phi\big(P_i, I^n_{\hat{j}}, \tau^n\big)$ at (P_i, I_j, τ^n). Similarly, expanding (D.19) at (P_i, I_j, τ^{n+1}) gives

$$\left[\Phi^{n+1}L\phi^n\right]_{i,j} = (\mathcal{L}\phi)^{n+1}_{i,j} + O(\Delta\tau + \Delta P_{\max} + (\Delta I_{\max})^2) \tag{D.22}$$

For convenience, we define

$$W\big(P_i, I_j, \tau^{n+1}; c\big) = \big(\phi_\tau + (c + a(c))\phi_I - (c - a(c))P - \mathcal{L}\phi\big)^{n+1}_{i,j}. \tag{D.23}$$

Substituting equations (D.17), (D.21-D.22) and $(\phi_I)^n_{i,j} = (\phi_I)^{n+1}_{i,j} + O(\Delta\tau)$ into $\mathcal{G}^{n+1}_{i,j}$ given in (3.2.1), then writing $\mathcal{G}^{n+1}_{i,j}$ in terms of notation (D.23), gives

$$\begin{aligned}&\mathcal{G}^{n+1}_{i,j}\big(h, \phi^{n+1}_{i,j}, \left\{\phi^{n+1}_{k,j}\right\}_{k\neq i}, \left\{\phi^n_{i,j}\right\}\big) \\ &= \inf_{(\zeta^n_{i,j}, \zeta^{n+1}_{i,j}) \in C^{n+1}_j} \Big\{(1-\theta)W\big(P_i, I_j, \tau^{n+1}; \zeta^{n+1}_{i,j}\big) + \theta W\big(P_i, I_j, \tau^{n+1}; \zeta^n_{i,j}\big) \\ &\quad + O\left(\Delta\tau + \Delta P_{\max} + (\Delta I_{\max})^2/\Delta\tau + (\Delta I_{\max})^2\right)\Big\} \\ &= \inf_{(\zeta^n_{i,j}, \zeta^{n+1}_{i,j}) \in C^{n+1}_j} \left\{(1-\theta)W\big(P_i, I_j, \tau^{n+1}; \zeta^{n+1}_{i,j}\big) + \theta W\big(P_i, I_j, \tau^{n+1}; \zeta^n_{i,j}\big) + O(h)\right\},\end{aligned} \tag{D.24}$$

where we use equations (2.3.1) to unify the mesh size/timestep size using h. Note that the multiplier function for the O-notation in equation (D.24) is a function of $\zeta^n_{i,j}$ and $\zeta^{n+1}_{i,j}$. In other words, we may write $O(h) = H\big(\zeta^n_{i,j}, \zeta^{n+1}_{i,j}\big)h$, where H is a bounded function of

$\zeta^n_{i,j}, \zeta^{n+1}_{i,j}$.

According to Lemma D.1, by taking $\Delta\tau$ sufficiently small, given equations (2.2.1), (2.2.3) and (2.2.5), we can relax the constraint $(\zeta^n_{i,j}, \zeta^{n+1}_{i,j}) \in C^{n+1}_j$ in above equation (D.24) to $\zeta^{n+1}_{i,j} \in C(I_j)$, $\zeta^n_{i,j} \in C\big(I^n_{\hat{j}}\big)$.

The above argument together with Lemma D.2 allows us to write equation (D.24) as

$$\begin{aligned}
&\mathcal{G}^{n+1}_{i,j}\big(h, \phi^{n+1}_{i,j}, \{\phi^{n+1}_{k,j}\}_{k\neq i}, \{\phi^n_{i,j}\}\big) \\
&= (1-\theta) \inf_{\zeta^{n+1}_{i,j}\in C(I_j)} W\big(P_i, I_j, \tau^{n+1}; \zeta^{n+1}_{i,j}\big) + \theta \inf_{\zeta^n_{i,j}\in C(I^n_{\hat{j}})} W\big(P_i, I_j, \tau^{n+1}; \zeta^n_{i,j}\big) + Q(h)h,
\end{aligned} \tag{D.25}$$

where Q is a bounded function satisfying (for some function $H\big(\zeta^n_{i,j}, \zeta^{n+1}_{i,j}\big)$)

$$Q(h) \leq \sup_{\zeta^n_{i,j}\in C(I^n_{\hat{j}}), \zeta^{n+1}_{i,j}\in C(I_j)} \big|H\big(\zeta^n_{i,j}, \zeta^{n+1}_{i,j}\big)\big|.$$

By (D.25) and (D.16), the left hand side of the equation (D.15) can be written using the notation (D.23) as

$$\begin{aligned}
&\lim_{h\to 0} \Big| \inf_{c\in C(I_j)} W\big(P_i, I_j, \tau^{n+1}; c\big) - (1-\theta) \inf_{\zeta^{n+1}_{i,j}\in C(I_j)} W\big(P_i, I_j, \tau^{n+1}; \zeta^{n+1}_{i,j}\big) \\
&\qquad -\theta \inf_{\zeta^n_{i,j}\in C(I^n_{\hat{j}})} W\big(P_i, I_j, \tau^{n+1}; \zeta^n_{i,j}\big) - Q(h)h \Big| \\
&= \lim_{h\to 0} \Big| \theta \inf_{c\in C(I_j)} W\big(P_i, I_j, \tau^{n+1}; c\big) - \theta \inf_{\zeta^n_{i,j}\in C(I^n_{\hat{j}})} W\big(P_i, I_j, \tau^{n+1}; \zeta^n_{i,j}\big) - Q(h)h \Big| \\
&= 0,
\end{aligned} \tag{D.26}$$

where the last equality follows because as $h \to 0$, then $\Delta\tau \to 0$, and hence $C\big(I^n_{\hat{j}}\big) \to C(I_j)$. Therefore, we have proved the pointwise consistency result (D.15) for the semi-Lagrangian discretization (3.2.1-3.2.2). □

After presenting Lemmas D.1, D.2 and D.3, in the following we prove Lemma 3.19.

Let $\hat{\mathbf{x}} = (P, I, \tau)$ and let $\mathbf{x} = (P_i, I_j, \tau^{n+1})$ be a discrete mesh node. First we consider

the region $\hat{\mathbf{x}} \in \Omega = [0, P_{\max}] \times [0, I_{\max}] \times (0, T]$. According to the definition in (3.1.39), $\bar{F} = F$ in Ω, where, from (3.1.1),

$$F\big(D^2\phi(\hat{\mathbf{x}}), D\phi(\hat{\mathbf{x}}), \phi(\hat{\mathbf{x}}), \hat{\mathbf{x}}\big) = \inf_{c \in C(I)} \{\phi_\tau + (c + a(c))\phi_I - (c - a(c))P - \mathcal{L}\phi\} \tag{D.27}$$

Since $F(M, p, g, y)(M = D^2V, p = DV, g = V, y = \hat{\mathbf{x}})$ is a continuous function of (M, p, g, y), then $\bar{F}_* = \bar{F}^* = F$ in Ω. Consequently, inequalities (3.2.5-3.2.6) reduce to the point-wise consistency condition:

$$\begin{aligned}
&\lim_{h \to 0} \Big| F\big(D^2\phi(\mathbf{x}), D\phi(\mathbf{x}), \phi(\mathbf{x}), \mathbf{x}\big) - \mathcal{G}_{i,j}^{n+1}\big(h, \phi_{i,j}^{n+1}, \{\phi_{k,j}^{n+1}\}_{k \neq i}, \{\phi_{i,j}^{n}\}\big) \Big| \\
&= \lim_{h \to 0} \Big| \inf_{c \in C(I_j)} \Big\{ \big(\phi_\tau + (c + a(c))\phi_I - (c - a(c))P - \mathcal{L}\phi\big)_{i,j}^{n+1} \Big\} - \\
&\qquad\qquad \mathcal{G}_{i,j}^{n+1}\big(h, \phi_{i,j}^{n+1}, \{\phi_{k,j}^{n+1}\}_{k \neq i}, \{\phi_{i,j}^{n}\}\big) \Big| \\
&= 0 \qquad \forall n + 1 \geq 1,
\end{aligned} \tag{D.28}$$

which has been proved in Lemma D.3.

Next we need to prove inequalities (3.2.5-3.2.6) at the boundary region $(P, I, \tau) \in [0, P_{\max}] \times [0, I_{\max}] \times \{0\}$. This can be done by following the same method described in Chapter 8. We omit the details here because the Dirichlet boundary condition for the problem here is much easier compared with degenerate boundary conditions for the problem we will see in Chapter 8.

Appendix E

Regime-Switching Model Calibration

In this appendix we illustrate the procedure of calibrating the models in Section 5.2 to the market gas futures prices and options on futures.

E.1 Calibration to Futures

In this section we describe the calibration of the regime-switching model (5.2.8-5.2.9) to the market futures prices; the calibration procedure of the MR model (5.2.1-5.2.2) is similar but simpler, and hence omitted here.

E.1.1 Futures Price Valuation

Let $F^k(P,t,T)$ denote the natural gas futures price in regime k, $k \in \{0,1\}$, at time t with delivery at T, while the gas spot price resides at P. Assuming the risk neutral natural gas spot price follows the regime-switching model (5.2.8-5.2.9), we can write $F^k(P,t,T)$ as the risk neutral expectation of the spot price at T

$$F^k(p,t,T) = E^{\mathbb{Q}}[P(T) \mid P(t) = p, m(t) = k], \tag{E.1}$$

where $m(t)$ is the two-state Markov chain given in (5.2.7), representing the regime in which the risk neutral gas spot price resides at t. From equation (E.1), F^k satisfies two

PDEs that are coupled with each other given by

$$F_t^k + \left[\alpha^k\left(K_0^k - P\right) + S^k(t)P\right]F_P^k + \frac{1}{2}(\sigma^k)^2P^2F_{PP}^k + \lambda^{k\to(1-k)}(F^{1-k} - F^k) = 0 \ , \ k \in \{0,1\} \tag{E.2}$$

with the boundary conditions

$$F^k(P,T,T) = P \ , \ k \in \{0,1\}. \tag{E.3}$$

The solution to PDEs (E.2) has the form

$$F^k(P,t,T) = a^k(t,T) + b^k(t,T)P, \tag{E.4}$$

where functions a, b are independent of P. Substituting equation (E.4) into equations (E.2-E.3) gives an ODE system

$$\begin{aligned} a_t^k + \lambda^{k\to(1-k)}(a^{1-k} - a^k) + \alpha^k K_0^k b^k = 0 \\ b_t^k - \left[\alpha^k - S^k(t) + \lambda^{k\to(1-k)}\right]b^k + \lambda^{k\to(1-k)}b^{1-k} = 0 \ , \ k \in \{0,1\} \end{aligned} \tag{E.5}$$

subject to the boundary conditions

$$a^k(T,T) = 0 \ \ ; \ \ b^k(T,T) = 1 \ , \ k \in \{0,1\}. \tag{E.6}$$

Remark E.1. *For the regime-switching model, equation (E.4) and the ODE system (E.5-E.6) imply that the futures prices $F^k(t,T), k \in \{0,1\}$ at time t when the gas spot price is known are independent of the spot price volatilities σ^0, σ^1. Similar observations indicate that the futures price is independent of spot price volatility for the MR model. Consequently, the volatility needs to be calibrated using financial instruments other than futures contracts; in this thesis, we choose options on futures (see Section E.2 for the detailed calibration procedure).*

E.1.2 Calibration Procedure

Let $\theta = \{\alpha^k, K_0^k, \beta_A^k, \beta_{SA}^k, A_0, SA_0, \lambda^{k\to(1-k)} \mid k \in \{0,1\}\}$ denote the set of parameters we need to obtain through calibrating to the futures price data.

We use a least squares approach for the calibration. For each observation day t, we need to determine the regime in which the risk neutral gas spot price lies. Following the approach in [27, 91], we treat the regime number as a latent variable and reveal its value through calibration. Specifically, we perform the calibration by solving the following optimization problem:

$$\begin{aligned} &\min_{\theta} \sum_{t} \sum_{T} \left(\hat{F}^{\hat{k}(t;\theta)}\big(P(t), t, T; \theta\big) - F(t,T) \right)^2 \ , \text{ where} \\ &\hat{k}(t;\theta) = \underset{k\in\{0,1\}}{\arg\min} \sum_{T} \left(\hat{F}^{k}\big(P(t), t, T; \theta\big) - F(t,T) \right)^2 , \end{aligned} \tag{E.7}$$

where $F(t,T)$ is the market futures price on the observation day t with maturity T; $\hat{F}^{\hat{k}(t;\theta)}(P(t), t, T; \theta)$ is the corresponding model implied futures price in regime $\hat{k}(t;\theta)$ calculated numerically from equations (E.4-E.6) using the market spot price $P(t)$ and the parameter set θ. In equation (E.7), the range of t consists of all the observation days in our sample data and that of T covers the thirteen consecutive delivery months starting two months after the month of t. We use the MATLAB nonlinear least squares optimization routine to solve the problem (E.7). The routine is not guaranteed to work since it assumes the object function is differentiable.

Here an optimal regime number $\hat{k}(t;\theta)$ is determined for each observation date t and for a given parameter set θ. Let $\hat{\theta}$ be the parameters returned by the calibration procedure. Then the regime $\hat{k}(t;\hat{\theta})$ is where the realized market gas price actually resides at time t, revealed from the calibration. Note that the time spent in a given regime, as determined by calibration, will depend on the $\mathbb{P}$ measure transition densities, not the $\mathbb{Q}$ measure densities.

Remark E.2. *In [78, 27], the commodity spot price serves as an unknown parameter, instead of an observed market value, and is estimated from the calibration process. This*

may help improve the calibration results.

In [27], the estimated spot price is assured to be consistent with the simultaneously calibrated model volatilities and correlations. Note that in our case the spot price is observed from the market and the latent variable $\hat{k}(t;\theta)$ does not have such a consistency requirement. This is because the transition densities under the $\mathbb{P}$ and $\mathbb{Q}$ measures are not the same.

E.2 Calibration to Options on Futures

In this section, we calibrate the volatility using market European call options on natural gas futures. We demonstrate the calibration procedure only for the regime-switching model (5.2.8-5.2.9). Similar but simpler procedures follow for the MR model.

E.2.1 Futures Option Valuation

Let $\bar{V}^k(F,t,T_v)$ denote the European call option value in regime k at time t with maturity at T_v, where F represents the price of the underlying futures contract at time t. Let K denote the strike price of the option. Let $F^k(t,T)$ represent the price of the underlying futures contract in regime k at time t with maturity at T, where T satisfies $T \geq T_v$.

In NYMEX, the trading of a European option ends on the business day immediately preceding the expiration of the underlying futures contract. As a result, we can assume $T_v = T$, and we will thus use T as the maturity for both an option and its underlying futures contract.

We can write $\bar{V}^k(F,t,T)$ in the form of the risk neutral expectation as

$$\begin{aligned}\bar{V}^k(f,t,T) &= e^{-r(T-t)}E^{\mathbb{Q}}\left[\left(F^{m(T)}(T,T)-K\right)\mathbf{1}_{F^{m(T)}(T,T)\geq K} \mid F^k(t,T)=f, m(t)=k\right] \\ &= e^{-r(T-t)}E^{\mathbb{Q}}\left[\left(P(T)-K\right)\mathbf{1}_{P(T)\geq K} \mid a^k(t,T)+b^k(t,T)P(t)=f, m(t)=k\right],\end{aligned} \tag{E.8}$$

where $\mathbf{1}_{x\geq y}$ is an indicator function that returns 1 if $x \geq y$, or 0 if $x < y$; the second equality above uses the fact that $F^{m(T)}(T,T) = P(T)$ at maturity T as well as the relation (E.4) between futures price F and spot price P at time t assuming the risk neutral gas spot price follows the regime-switching model (5.2.8-5.2.9). Let $V^k(P,t,T)$ represent a synthetic European call option on spot price P at time t, in regime k with maturity T and strike K. Then we can write $V^k(P,t,T)$ in the form of the risk neutral expectation as

$$V^k(p,t,T) = e^{-r(T-t)} E^{\mathbb{Q}} \left[\big(P(T) - K\big) \mathbf{1}_{P(T)\geq K} \mid P(t) = p, m(t) = k \right]. \tag{E.9}$$

Comparing equations (E.8) and (E.9), we have

$$V^k \left(\frac{f - a^k(t,T)}{b^k(t,T)}, t, T \right) = \bar{V}^k(f,t,T), \tag{E.10}$$

where a^k and b^k are computed from the ODE system (E.5-E.6). As a result, we can compute $\bar{V}^k(F,t,T)$ using equation (E.10) as long as we are able to solve for $V^k(P,t,T)$.

Let r denote the constant riskless interest rate. Assuming that the spot price process follows SDE (5.2.8-5.2.9) and using the risk neutral expectation formulation (E.9), we find that the synthetic option value V^k satisfies the coupled PDEs

$$\begin{aligned} V^k_t + \big[\alpha^k\big(K^k_0 - P\big) + S^k(t)P\big] V^k_P + \frac{1}{2}(\sigma^k)^2 P^2 V^k_{PP} - rV^k + \\ \lambda^{k\to(1-k)}(V^{1-k} - V^k) = 0 \ , \ k \in \{0,1\} \end{aligned} \tag{E.11}$$

subject to the boundary conditions

$$V^k(P,T,T) = \max[P - K, 0] \ , \ k \in \{0,1\}, \tag{E.12}$$

We will solve equations (E.11-E.12) numerically in the computational domain $P \in [0, P_{\max}]$ with $P_{\max} \gg K$. For this purpose, we need to impose boundary conditions at the computational boundary $P = 0$ and $P = P_{\max}$. At the $P = 0$ boundary, taking the

limit as $P \to 0$, equations (E.11) become

$$V_t^k + \alpha^k K_0^k V_P^k - rV^k + \lambda^{k\to(1-k)}(V^{1-k} - V^k) = 0 \ , \ k \in \{0,1\}. \tag{E.13}$$

Since $\alpha^k K_0^k \geq 0$ for all variations of the regime-switching model we consider (see Section 5.2.2), the characteristics are outgoing or zero in the P direction at $P = 0$. As a result, we can solve equations (E.13) without requiring additional boundary conditions, as we do not need information from outside the computational domain. At the $P = P_{\max}$ boundary, we make the assumption that $V^k(P_{\max}, t, T) \to$ payoff. In other words, we impose the Dirichlet boundary condition

$$V^k(P_{\max}, t, T) = P_{\max} - K \ , \ k \in \{0,1\}. \tag{E.14}$$

The error introduced by this approximation will be small if $P_{\max}$ is sufficiently large.

During calibration, we will solve equations (E.11-E.14) using a standard fully implicit finite difference scheme that is stable, monotone and consistent and hence the scheme converges to the unique solution of equations (E.11-E.14). We skip the details here.

E.2.2 Calibration Procedure

The first step of the calibration is to determine the regime in which the underlying futures contract lies; the approach is given in Section E.1.2. After determining the optimal regime, denoted by $\hat{k}$, for the futures contract at time t, we use a least squares approach to calibrate the volatility σ^k by solving

$$\min_{\sigma^0,\sigma^1} \sum_K \left(\bar{V}^{\hat{k}}\big(F(t,T), t, T; \theta, K, \sigma^0, \sigma^1\big) - C(t,T;K) \right)^2, \tag{E.15}$$

where $C(t,T;K)$ represents the value of a market call option on futures at time t with maturity at T and strike price at K; $\bar{V}^{\hat{k}}$ is the corresponding model implied option value in regime $\hat{k}$ using the market futures price $F(t,T)$, the parameter set $\theta = \{\alpha^k, K_0^k, \beta_A^k, \beta_{SA}^k, A_0, SA_0, \lambda^{k\to(1-k)} \,|\, k \in \{0,1\}\}$ and the volatility pair σ^0, σ^1. The values of $\bar{V}^{\hat{k}}$ are computed

by first solving PDEs (E.11-E.14) numerically to obtain values of V^k at time t and at discrete grid nodes in the P direction and then linearly interpolating these discrete values using equation (E.10). We choose the mesh size so that the error in $\bar{V}^{\hat{k}}$ is less than 10^{-2}.

Appendix F

Proofs for Discrete Withdrawal GMWB Variable Annuities

In this appendix we give the proofs for Lemma 7.2, Proposition 7.3, Proposition 7.6 and Theorem 7.13.

F.1 Proof for Lemma 7.2

The proof of Lemma 7.2 relies on the following Lemma.

Lemma F.1. *This Lemma provides some useful results. Given any $a, b \in \mathbb{R}$, it is straightforward to verify that*

$$|\max(a,0) - \max(b,0)| \leq |a-b|. \tag{F.1}$$

Suppose $X(x), Y(x)$ are functions defined for some bounded compact domain $x \in D$, then according to [46], we have

$$\left|\sup_{x\in D} X(x) - \sup_{y\in D} Y(y)\right| \leq \sup_{x\in D} |X(x) - Y(x)|. \tag{F.2}$$

After presenting Lemma F.1, in the following we prove Lemma 7.2. At first, from equation (7.2.6), we know $V(W, A, \tau_O^{k+})$ exists for all $(W, A) \in [0, W_{\max}] \times [0, w_0]$. To

prove the uniform continuity of $V(W, A, \tau_O^{k+})$ on (W, A), by definition, we need to show that $\forall \epsilon > 0$, $\exists \sigma > 0$, such that $\forall (W', A'), (W'', A'') \in [0, W_{\max}] \times [0, w_0]$ satisfying $\sqrt{(W' - W'')^2 + (A' - A'')^2} < \sigma$, we have $|V(W', A', \tau_O^{k+}) - V(W'', A'', \tau_O^{k+})| < \epsilon$.

Let $Y(\gamma; W, A)$ be a function of $\gamma \in [0, A]$ defined as

$$Y(\gamma; W, A) = V(\max(W - \gamma, 0), A - \gamma, \tau_O^k) + f(\gamma), \tag{F.3}$$

where $f(\gamma)$ is given in (7.2.5). Without loss of generality, we assume $A' \geq A''$. We can write

$$\begin{aligned}
&|V(W', A', \tau_O^{k+}) - V(W'', A'', \tau_O^{k+})| \\
&\leq \left| V(W', A', \tau_O^{k+}) - \sup_{\gamma^k \in [0, A'']} \left[V\big(\max(W' - \gamma^k, 0), A' - \gamma^k, \tau_O^k\big) + f(\gamma^k) \right] \right| \\
&\quad + \left| \sup_{\gamma^k \in [0, A'']} \left[V\big(\max(W' - \gamma^k, 0), A' - \gamma^k, \tau_O^k\big) + f(\gamma^k) \right] - V(W'', A'', \tau_O^{k+}) \right| \\
&\leq \left[\sup_{\gamma^k \in [0, A']} \left\{ Y(\gamma^k; W', A') \right\} - \sup_{\gamma^k \in [0, A'']} \left\{ Y(\gamma^k; W', A') \right\} \right] \\
&\quad + \sup_{\gamma^k \in [0, A'']} \left| V\big(\max(W' - \gamma^k, 0), A' - \gamma^k, \tau_O^k\big) - V\big(\max(W'' - \gamma^k, 0), A'' - \gamma^k, \tau_O^k\big) \right|,
\end{aligned} \tag{F.4}$$

where the term inside the bracket of the last inequality above is due to the definition in (F.3)[1], and the last term in the last inequality above is due to (F.2). Next we will consider these two expressions individually.

Let us first consider the expression $\sup_{\gamma^k \in [0, A']} Y(\gamma^k; W', A') - \sup_{\gamma^k \in [0, A'']} Y(\gamma^k; W', A')$.

[1] This term is always positive since $[0, A''] \subseteq [0, A']$ and the functions in the two sup expressions are identical. Thus there is no need to take the absolute value for this term.

We can write

$$\begin{aligned}
&\sup_{\gamma^k\in[0,A']} Y(\gamma^k;W',A') - \sup_{\gamma^k\in[0,A'']} Y(\gamma^k;W',A') \\
&= \max\left[\sup_{\gamma^k\in[0,A'']} Y(\gamma^k;W',A'), \sup_{\gamma^k\in(A'',A']} Y(\gamma^k;W',A')\right] - \sup_{\gamma^k\in[0,A'']} Y(\gamma^k;W',A') \\
&= \max\left[0, \sup_{\gamma^k\in(A'',A']} Y(\gamma^k;W',A') - \sup_{\gamma^k\in[0,A'']} Y(\gamma^k;W',A')\right]
\end{aligned} \tag{F.5}$$

Since $V(W,A,\tau_O^k)$ is uniformly continuous on (W,A) and since $f(\gamma^k)$ is uniformly continuous, we obtain

$$\lim_{\gamma^k\to[\gamma]^-} Y(\gamma^k;W',A') = \lim_{\gamma^k\to[\gamma]^+} Y(\gamma^k;W',A'), \quad \forall\gamma\in[0,A']. \tag{F.6}$$

According to (F.6), we have

$$\begin{aligned}
\lim_{\gamma^k\to[A'']^+} Y(\gamma^k;W',A') &= \lim_{\gamma^k\to[A'']^-} Y(\gamma^k;W',A') \\
&\le \sup_{\gamma^k\in[0,A'']} Y(\gamma^k;W',A').
\end{aligned} \tag{F.7}$$

Since we have

$$\lim_{A'\to A''} \sup_{\gamma^k\in(A'',A']} Y(\gamma^k;W',A') = \lim_{\gamma^k\to[A'']^+} Y(\gamma^k;W',A'), \tag{F.8}$$

(F.7-F.8) imply that

$$\lim_{A'\to A''} \sup_{\gamma^k\in(A'',A']} Y(\gamma^k;W',A') \le \sup_{\gamma^k\in[0,A'']} Y(\gamma^k;W',A'). \tag{F.9}$$

This together with (F.5) shows that, $\forall\epsilon>0$,

$$\exists\sigma_0>0,\ \forall|A'-A''|<\sigma_0, \quad \sup_{\gamma^k\in[0,A']} Y(\gamma^k;W',A') - \sup_{\gamma^k\in[0,A'']} Y(\gamma^k;W',A') < \epsilon/2. \tag{F.10}$$

Let us now consider the last term in inequality (F.4). Since $V(W,A,\tau_O^k)$ is uniformly

continuous on (W, A), then $\forall \epsilon > 0$, $\exists \sigma_1 > 0$, $\forall (W_1, A_1), (W_2, A_2) \in [0, W_{\max}] \times [0, w_0]$ satisfying $\sqrt{(W_1 - W_2)^2 + (A_1 - A_2)^2} < \sigma_1$, we have

$$\left| V\left(W_1, A_1, \tau_O^k\right) - V\left(W_2, A_2, \tau_O^k\right) \right| < \epsilon/2. \tag{F.11}$$

Let

$$W_1 = \max(W' - \gamma^k, 0),\ W_2 = \max(W'' - \gamma^k, 0),\ A_1 = A' - \gamma^k,\ A_2 = A'' - \gamma^k. \tag{F.12}$$

Inequality (F.1) implies that $\sqrt{(W_1 - W_2)^2 + (A_1 - A_2)^2} \leq \sqrt{(W' - W'')^2 + (A' - A'')^2}$. Consequently, if $(W', A'), (W'', A'')$ satisfy

$$\sqrt{(W' - W'')^2 + (A' - A'')^2} < \sigma_1, \tag{F.13}$$

then (F.11-F.12) lead to

$$\left| V\left(\max(W' - \gamma^k, 0), A' - \gamma^k, \tau_O^k\right) - V\left(\max(W'' - \gamma^k, 0), A'' - \gamma^k, \tau_O^k\right) \right| < \epsilon/2. \tag{F.14}$$

As a result, $\forall \epsilon > 0$, $\exists \sigma_0, \sigma_1$, $\forall (W', A'), (W'', A'') \in [0, W_{\max}] \times [0, w_0]$ satisfying $\sqrt{(W' - W'')^2 + (A' - A'')^2} < \min(\sigma_0, \sigma_1)$, then according to (F.4), (F.10) and (F.13-F.14), we obtain

$$|V(W', A', \tau_O^{k+}) - V(W'', A'', \tau_O^{k+})| < \epsilon. \tag{F.15}$$

This verifies the uniform continuity of function $V(W, A, \tau_O^{k+})$ on $(W, A) \in [0, W_{\max}] \times [0, w_0]$ by definition.

F.2 Proof for Proposition 7.3

In this section we prove Proposition 7.3.

The terminal boundary condition (7.2.9) implies that $V(W, A, \tau_O^0)$ is uniformly continuous on $(W, A) \in [0, W_{\max}] \times [0, w_0]$. Then according to Lemma 7.2, $V(W, A, \tau_O^{0+})$

is a uniformly continuous function of (W, A). Since in addition the boundary equation (7.2.12) at $W = 0$ is the limit of equation (7.2.7) towards the boundary and boundary equation (7.2.13) at $W = W_{\max}$ is a standard Dirichlet condition, then there exists a unique continuous classical solution (and hence a viscosity solution) to equation (7.2.7) in the domain $\bar{\Omega}_0$ with initial condition $V(W, A, \tau_O^{0+})$ and boundary conditions (7.2.12-7.2.13). Consequently, the proposition holds by applying the above arguments to each interval $[\tau_O^{k+}, \tau_O^{k+1}]$, $k = 1, \ldots, K-1$.

F.3 Proof for Proposition 7.6

In this section we prove Proposition 7.6.

Let us define $g(\gamma_{i,j}^n)$ as a piecewise linear function of $\gamma_{i,j}^n \in [0, A_j]$ constructed using the discrete values from the set $\left\{ \phi_{\hat{i},\hat{j}}^n \;\middle|\; \gamma_{i,j}^n \in \mathcal{A}_j \right\}$ by linear interpolation. Without loss of generality, we assume $A_j > W_i > G^k$. Since $f(\gamma_{i,j}^n)$ is a piecewise linear function on $\gamma \in [0, G^k]$ and $\gamma \in (G^k, A_j]$, and since all the nodes $0, G^k, W_i, A_j$ belong to $\mathcal{A}_j$, then the supremum of function $g(\gamma_{i,j}^n) + f(\gamma_{i,j}^n)$ occurs at a node $\gamma_{i,j}^n \in \mathcal{A}_j$. Consequently, we have

$$\sup_{\gamma_{i,j}^n \in [0, A_j]} \left[g(\gamma_{i,j}^n) + f(\gamma_{i,j}^n) \right] = \sup_{\gamma_{i,j}^n \in \mathcal{A}_j} \left[\phi_{\hat{i},\hat{j}}^n + f\big(\gamma_{i,j}^n\big) \right]. \tag{F.16}$$

For any $\gamma_{i,j}^n \in \mathcal{A}_j$, equations (7.3.5) and (7.3.1) imply $g(\gamma_{i,j}^n) = \phi\big(\max(W_i - \gamma_{i,j}^n, 0), A_j - \gamma_{i,j}^n, \tau^n\big) + O(h^2)$. Since in addition ϕ is smooth and the distance between any two consecutive elements in $\mathcal{A}_j$ is bounded by $O(h)$, we have (recall that $W_i \in \mathcal{A}_j$, so that $\max(W_i - \gamma_{i,j}^n, 0)$ is smooth between the nodes in $\mathcal{A}_j$)

$$g(\gamma_{i,j}^n) = \phi\big(\max(W_i - \gamma_{i,j}^n, 0), A_j - \gamma_{i,j}^n, \tau^n\big) + O(h^2), \quad \forall \gamma_{i,j}^n \in [0, A_j]. \tag{F.17}$$

Equation (F.17) and a result similar to Lemma D.2 imply that

$$\sup_{\gamma_{i,j}^n \in [0, A_j]} \left[g(\gamma_{i,j}^n) + f(\gamma_{i,j}^n) \right] = \sup_{\gamma_{i,j}^n \in [0, A_j]} \left[\phi\big(\max(W_i - \gamma_{i,j}^n, 0), A_j - \gamma_{i,j}^n, \tau^n\big) + f\big(\gamma_{i,j}^n\big) \right] + O(h^2). \tag{F.18}$$

Therefore, (7.4.4) follows from (F.16) and (F.18).

Finally, equation (7.4.3) holds according to (7.4.4) and the following equation implied from (7.3.5) and (7.3.1):

$$\sup_{\gamma^n_{i,j}\in[0,A_j]}\Big[\phi^n_{i,\hat{j}}+f\big(\gamma^n_{i,j}\big)\Big]=\sup_{\gamma^n_{i,j}\in[0,A_j]}\big[\phi\big(\max(W_i-\gamma^n_{i,j},0),A_j-\gamma^n_{i,j},\tau^n\big)+f\big(\gamma^n_{i,j}\big)\big]+O(h^2). \tag{F.19}$$

F.4 Proof for Theorem 7.13

In this section we prove Theorem 7.13.

Let $V^h(\tau_O^{0+})$ denote the approximate solution computed by (7.3.6) at τ_O^{0+} with the mesh size/timestep parameter h. $V^h(\tau_O^{0+})$ is only defined at mesh nodes (W_i, A_j). Let $V_I^h(\tau_O^{0+})$ denote the value of the approximate solution which is interpolated using linear interpolation for any point (W, A). Let $V(\tau_O^{0+})$ be the exact solution to equation (7.2.6). Here we suppress the variables (W, A) in the above notation. Since (7.3.6) is a consistent discretization of equation (7.2.6), then $V_I^h(\tau_O^{0+})$ converges to $V(\tau_O^{0+})$ as $h \to 0$.

Let any $\mathbf{x} = (W, A, \tau) \in \bar{\Omega}_0$, where $\bar{\Omega}_0 = [0, W_{\max}] \times [0, w_0] \times [\tau_O^{0+}, \tau_O^1]$ as defined in (7.2.14). Let $V^h\big(V_I^h(\tau_O^{0+})\big)$ denote the approximate solution resulting from equation (7.3.7) with initial condition $V_I^h(\tau_O^{0+})$ at mesh nodes $(W_i, A_j, \tau^{n+1}) \in \bar{\Omega}_0$. Accordingly, let $V_I^h\big(\mathbf{x}, V_I^h(\tau_O^{0+})\big)$ be the value of the approximate solution at $\mathbf{x}$ obtained by linear interpolation using $V^h\big(V_I^h(\tau_O^{0+})\big)$ defined only at mesh nodes. Let $V\big(\mathbf{x}; V(\tau_O^{0+})\big)$ and $V\big(\mathbf{x}; V_I^h(\tau_O^{0+})\big)$ denote the unique viscosity solution to equation (7.2.7) and boundary conditions (7.2.12-7.2.13), with initial condition $V(\tau_O^{0+})$ and $V_I^h(\tau_O^{0+})$, respectively. Since $V_I^h(\tau_O^{0+}) \to V(\tau_O^{0+})$ as $h \to 0$, we have

$$V\Big(\mathbf{x}; V_I^h\big(\tau_O^{0+}\big)\Big) \to V\Big(\mathbf{x}; V\big(\tau_O^{0+}\big)\Big) \quad \text{as } h \to 0. \tag{F.20}$$

According to Lemmas 7.8, 7.10 and 7.12, scheme (7.3.7) is l_∞ stable and monotone, and pointwise consistent to PDE (7.2.7) and its boundary conditions (7.2.12-7.2.13). Thus,

convergence results in [11, 6] imply that

$$V_I^h\Big(\mathbf{x}; V_I^h\big(\tau_O^{0+}\big)\Big) \to V\Big(\mathbf{x}; V_I^h\big(\tau_O^{0+}\big)\Big) \quad \text{as } h \to 0. \tag{F.21}$$

Using equation (F.20-F.21), we have

$$\begin{aligned}
&\Big|V_I^h\Big(\mathbf{x}; V_I^h\big(\tau_O^{0+}\big)\Big) - V\Big(\mathbf{x}; V\big(\tau_O^{0+}\big)\Big)\Big| \\
&\leq \Big|V_I^h\Big(\mathbf{x}; V_I^h\big(\tau_O^{0+}\big)\Big) - V\Big(\mathbf{x}; V_I^h\big(\tau_O^{0+}\big)\Big)\Big| + \Big|V\Big(\mathbf{x}; V_I^h\big(\tau_O^{0+}\big)\Big) - V\Big(\mathbf{x}; V\big(\tau_O^{0+}\big)\Big)\Big| \qquad \text{(F.22)} \\
&\to 0 \qquad \text{as } h \to 0.
\end{aligned}$$

Thus we prove the Theorem in $\bar{\Omega}_0$. Equation (F.22) implies that $V_I^h(\tau_O^{1+}) \to V(\tau_O^{1+})$ as $h \to 0$. Consequently, the Theorem follows by sequentially applying the above argument to regions $\bar{\Omega}_k$, $k = 1, \ldots, K-1$.

Appendix G

Derivation of GMWB Variable Annuity Pricing Equation under the Continuous Withdrawal Scenario

This appendix heuristically derives the GMWB variable annuity pricing PDE from the impulse control formulation (8.2.17) using dynamic programming (Bellman's Principle) and Ito's Lemma. The derivation is based on the arguments in [60]. In the following we assume $\hat{V}$ is smooth.

First, let us consider all the paths $(W(s), A(s))$ generated by optimal controls and the random factor dZ, where the impulse operation (i.e., withdraw a finite amount γ^k instantaneously) never occurs during the interval $[t, T]$. (i.e., it is not optimal to have a control (γ^1, t_s^1) with $t \leq t_s^1 \leq T$). In this case, the impulse control formulation (8.2.17) reduces to

$$\hat{V}(W, A, t-) = \sup_{\hat{\gamma}(s) \in [0, G_r]} E^{\mathbb{Q}} \left[\int_t^T e^{-r(s-t)} \hat{\gamma}(s) ds + e^{-r(T-t)} \hat{V}\big(W(T), A(T), T\big) \right]. \quad \text{(G.1)}$$

This is a regular bounded stochastic control problem with $\hat{\gamma}$ as the only control variable. As such, we can use an approach similar to that in Appendix A to derive the following HJB equation from (G.1), (8.2.12) and (8.2.14) (assuming all assumptions in Appendix A

are satisfied)

$$\hat{V}_t + \mathcal{L}\hat{V} + \sup_{\hat{\gamma}\in[0,G_r]} \big(\hat{\gamma} - \hat{\gamma}\hat{V}_W - \hat{\gamma}\hat{V}_A\big) = 0, \quad \text{if } t_s^1 \in [t,T] \text{ does not exist,} \tag{G.2}$$

where the operator $\mathcal{L}$ is given in (7.2.8). In the following, we only consider the cases where the first optimal impulse time $t_s^1 \le T$.

For an arbitrary path $(W(s), A(s))$ with $t_s^1 \le T$, starting from $(W(t_s^1-), A(t_s^1-), t_s^1-)$, the instant infinitesimally before the impulse time t_s^1, equation (8.2.17) still follows. Consequently, from (8.2.17) we have

$$\begin{aligned}
&\hat{V}\big(W(t_s^1-), A(t_s^1-), t_s^1-\big) \\
&= \sup_{\substack{\hat{\gamma}(s)\in[0,G_r]\\ \gamma^k\in[0,A(t_s^{k-})]\\ t_s^k\in[t,T]}} E^{\mathbb{Q}}\Bigg[\int_{t_s^1}^{T} e^{-r(s-t_s^1)}\hat{\gamma}(s)ds + \sum_{k=1} e^{-r(t_s^k-t_s^1)}\Big[(1-\kappa)\gamma^k - c\Big] \\
&\qquad\qquad + e^{-r(T-t_s^1)}\hat{V}\big(W(T), A(T), T\big)\Bigg] \\
&= \sup_{\substack{\hat{\gamma}(s)\in[0,G_r]\\ \gamma^k\in[0,A(t_s^{k-})]\\ t_s^k\in[t,T]}} E^{\mathbb{Q}}\Bigg[(1-\kappa)\gamma^1 - c + \Bigg[\int_{t_s^1}^{T} e^{-r(s-t_s^1)}\hat{\gamma}(s)ds + \sum_{k=2} e^{-r(t_s^k-t_s^1)}\Big[(1-\kappa)\gamma^k - c\Big] \\
&\qquad\qquad + e^{-r(T-t_s^1)}\hat{V}\big(W(T), A(T), T\big)\Bigg]\Bigg] \\
&= \sup_{\gamma^1\in[0,A(t_s^{1-})]} \Big[(1-\kappa)\gamma^1 - c + \hat{V}\big(W(t_s^1), A(t_s^1), t_s^1\big)\Big] \qquad \text{(G.3)} \\
&= \sup_{\gamma^1\in[0,A(t_s^{1-})]} \Big[(1-\kappa)\gamma^1 - c + \hat{V}\big(\max(W(t_s^1-)-\gamma^1, 0), A(t_s^1-)-\gamma^1, t_s^1\big)\Big], \qquad \text{(G.4)}
\end{aligned}$$

where the expectation $E^{\mathbb{Q}}$ is conditional on the known value $(W(t_s^1-), A(t_s^1-), t_s^1-)$; equation (G.3) above follows from dynamic programming for the impulse control problems [60, 16]; equation (G.4) above is due to dynamics (8.2.13) and (8.2.15).

In particular, if $t_s^1 = t$, that is, if the first impulse operation occurs at time t, equation

(G.4) reduces to

$$\begin{aligned}
&\hat{V}\big(W(t-),A(t-),t\big)\\
&=\hat{V}\big(W,A,t\big)=\sup_{\gamma\in(0,A]}\Big[\hat{V}(\max(W-\gamma,0),A-\gamma,t)+(1-\kappa)\gamma-c\Big],\quad \text{if } t=t_s^1,
\end{aligned}\tag{G.5}$$

where $\hat{V}\big(W(t-),A(t-),t\big)=\hat{V}\big(W,A,t\big)$ since we have assumed $\hat{V}$ is continuous on t and used the notation $(W(t-),A(t-))=(W,A)$ as the starting point of the path $W(s),A(s)$. Note that we can require $\gamma>0$ in (G.5) because $\gamma=0$ is never optimal due to the positive fixed cost c.

Next, assuming $t\le t_s^1\le T$, we can rewrite equation (8.2.17) as

$$\begin{aligned}
&\hat{V}(W,A,t-) \qquad\qquad \text{if } t\le t_s^1\le T\\
&=\sup_{\substack{\hat{\gamma}(s)\in[0,G_r]\\ \gamma^k\in[0,A(t_s^{k-})]\\ t_s^k\in[t,T]}} E^{\mathbb{Q}}\Bigg[\int_t^{t_s^1}e^{-r(s-t)}\hat{\gamma}(s)ds+\int_{t_s^1}^{T}e^{-r(s-t)}\hat{\gamma}(s)ds+e^{-r(t_s^1-t)}\Big[(1-\kappa)\gamma^1-c\Big]\\
&\qquad\qquad +\sum_{k=2}e^{-r(t_s^k-t)}\Big[(1-\kappa)\gamma^k-c\Big]+e^{-r(T-t)}\hat{V}\big(W(T),A(T),T\big)\Bigg]\\
&=\sup_{\substack{\hat{\gamma}(s)\in[0,G_r]\\ \gamma^k\in[0,A(t_s^{k-})]\\ t_s^k\in[t,T]}} E^{\mathbb{Q}}\Bigg[\int_t^{t_s^1}e^{-r(s-t)}\hat{\gamma}(s)ds+e^{-r(t_s^1-t)}\Big[(1-\kappa)\gamma^1-c\Big]\\
&\qquad\qquad +e^{-r(t_s^1-t)}\Bigg[\int_{t_s^1}^{T}e^{-r(s-t_s^1)}\hat{\gamma}(s)ds+\sum_{k=2}e^{-r(t_s^k-t_s^1)}\Big[(1-\kappa)\gamma^k-c\Big]\\
&\qquad\qquad +e^{-r(T-t_s^1)}\hat{V}\big(W(T),A(T),T\big)\Bigg]\Bigg]\\
&=\sup_{\substack{\hat{\gamma}(s)\in[0,G_r]\\ \gamma^1\in[0,A(t_s^{1-})]\\ t_s^1\in[t,T]}} E^{\mathbb{Q}}\Bigg[\int_t^{t_s^1}e^{-r(s-t)}\hat{\gamma}(s)ds+e^{-r(t_s^1-t)}\Big[(1-\kappa)\gamma^1-c\Big]+e^{-r(t_s^1-t)}\hat{V}\big(W(t_s^1),A(t_s^1),t_s^1\big)\Bigg],
\end{aligned}\tag{G.6}$$

where the last equality again follows from dynamic programming. The last equation in

(G.6) implies that

$$\begin{aligned}&\hat{V}\big(W,A,t\big)\\&=\hat{V}\big(W,A,t-\big)\geq\sup_{\gamma\in(0,A]}\Big[\hat{V}(\max(W-\gamma,0),A-\gamma,t)+(1-\kappa)\gamma-c\Big],\quad\text{if } t\leq t_s^1,\end{aligned}\tag{G.7}$$

where the inequality above is obtained by substituting $t_s^1 = t$ into (G.6) and using the assumption that the supremum in (G.6) is achieve at the stopping times equal to or greater than t. Following an argument similar to the above, we can show that (G.7) also holds if the impulse operation never occurs.

Let us now study equation (G.6) from another perspective. Assuming the first optimal impulse time $t \leq t_s^1 \leq T$, the last equality in (G.6) implies that

$$\begin{aligned}\hat{V}(W,A,t-)\geq&\sup_{\substack{\hat{\gamma}(s)\in[0,G_r]\\t_s^1\in(t,T]}}\sup_{\gamma^1\in[0,A(t_s^1-)]}E^{\mathbb{Q}}\Bigg[\int_t^{t_s^1}e^{-r(s-t)}\hat{\gamma}(s)ds\qquad\qquad(\text{G.8})\\&+e^{-r(t_s^1-t)}\Big[(1-\kappa)\gamma^1-c+\hat{V}\big(W(t_s^1),A(t_s^1),t_s^1\big)\Big]\Bigg]\\=&\sup_{\substack{\hat{\gamma}(s)\in[0,G_r]\\t_s^1\in(t,T]}}E^{\mathbb{Q}}\left[\int_t^{t_s^1}e^{-r(s-t)}\hat{\gamma}(s)ds+e^{-r(t_s^1-t)}\hat{V}\big(W(t_s^1-),A(t_s^1-),t_s^1-\big)\right].\qquad(\text{G.9})\end{aligned}$$

Note that we require $t_s^1 > t$ in the inequalities above, and the inequality (G.8) still holds in this case. In order to obtain (G.9) above, we first interchange operators $\sup_{\gamma^1\in[0,A(t_s^1-)]}$ and $E^{\mathbb{Q}}$ in (G.8) (we assume they are interchangeable) and then substitute equation (G.4).

Since by assumption t_s^1 is the first time for the impulse operation, then there exist only continuous withdrawals during the interval $[t, t_s^1-]$. Consequently, inequality (G.9) only consists of a regular stochastic control $\hat{\gamma}$. Let u be any fixed time satisfying $t < u \leq T$ and assume there is no impulse operation during $[t, u-]$. Then inequality (G.9) implies

that

$$\hat{V}(W,A,t) = \hat{V}(W,A,t-) \geq \sup_{\hat{\gamma}(s)\in[0,G_r]} E^{\mathbb{Q}}\left[\int_t^u e^{-r(s-t)}\hat{\gamma}(s)ds + e^{-r(u-t)}\hat{V}\big(W(u-),A(u-),u-\big)\right]. \tag{G.10}$$

Since inequality (G.10) contains only a regular stochastic control, we can use a procedure similar to that in Appendix A to derive the following inequality :

$$\hat{V}_t + \mathcal{L}\hat{V} + \sup_{\hat{\gamma}\in[0,G_r]}\big(\hat{\gamma} - \hat{\gamma}\hat{V}_W - \hat{\gamma}\hat{V}_A\big) \leq 0. \tag{G.11}$$

Let us look at equality (G.6) again. Suppose there exist an optimal pair (t_s^{1*}, γ^{1*}) with $t < t_s^{1*} \leq T$ that achieves the supremum in (G.6). As a result, equation (G.6) becomes

$$\begin{aligned}
&\hat{V}(W,A,t)\\
&= \hat{V}(W,A,t-)\\
&= \sup_{\hat{\gamma}(s)\in[0,G_r]} E^{\mathbb{Q}}\left[\int_t^{t_s^{1*}} e^{-r(s-t)}\hat{\gamma}(s)ds + e^{-r(t_s^{1*}-t)}\Big[(1-\kappa)\gamma^{1*} - c + \hat{V}\big(W(t_s^{1*}),A(t_s^{1*}),t_s^{1*}\big)\Big]\right]\\
&= \sup_{\hat{\gamma}(s)\in[0,G_r]} E^{\mathbb{Q}}\left[\int_t^{t_s^{1*}} e^{-r(s-t)}\hat{\gamma}(s)ds + e^{-r(t_s^{1*}-t)}\hat{V}\big(W(t_s^{1*}-),A(t_s^{1*}-),t_s^{1*}-\big)\right],
\end{aligned} \tag{G.12}$$

where the last equality is due to (G.4). Since there is no impulse control in (G.12), we can follow the steps in Appendix A to derive the following equation:

$$\hat{V}_t + \mathcal{L}\hat{V} + \sup_{\hat{\gamma}\in[0,G_r]}\big(\hat{\gamma} - \hat{\gamma}\hat{V}_W - \hat{\gamma}\hat{V}_A\big) = 0, \quad \text{if } t < t_s^{1*} \leq T. \tag{G.13}$$

From the discussions above, if the impulse operation never occurs, then equation (G.2) follows; if the first optimal impulse operation occurs at $t < t_s^1 \leq T$, then equation (G.13) follows; if $t_s^1 = t$, then equation (G.5) holds. This together with inequalities (G.7) and

(G.11) imply that

$$\begin{aligned}\hat{V}\big(W,A,t\big)-\sup_{\gamma\in(0,A]}\Big[\hat{V}(\max(W-\gamma,0),A-\gamma,t)+(1-\kappa)\gamma-c\Big]\geq 0\\ -\hat{V}_t-\mathcal{L}\hat{V}-\sup_{\hat{\gamma}\in[0,G_r]}\big(\hat{\gamma}-\hat{\gamma}\hat{V}_W-\hat{\gamma}\hat{V}_A\big)\geq 0\end{aligned}\tag{G.14}$$

and at least one of the inequalities above holds with equality. This is equivalent to the following HJB variational inequality

$$\begin{aligned}\min\Bigg\{&-\hat{V}_t-\mathcal{L}\hat{V}-\sup_{\hat{\gamma}\in[0,G_r]}\big(\hat{\gamma}-\hat{\gamma}\hat{V}_W-\hat{\gamma}\hat{V}_A\big),\\ &\hat{V}\big(W,A,t\big)-\sup_{\gamma\in(0,A]}\Big[\hat{V}(\max(W-\gamma,0),A-\gamma,t)+(1-\kappa)\gamma-c\Big]\Bigg\}=0.\end{aligned}\tag{G.15}$$

Using the notation $V(W,A,\tau)$ with $V(W,A,\tau)=\hat{V}(W,A,T-\tau)=\hat{V}(W,A,t)$, we can rewrite pricing equation (G.15) as

$$\begin{aligned}\min\Bigg\{&V_\tau-\mathcal{L}V-\sup_{\hat{\gamma}\in[0,G_r]}\big(\hat{\gamma}-\hat{\gamma}V_W-\hat{\gamma}V_A\big),\\ &V\big(W,A,t\big)-\sup_{\gamma\in(0,A]}\Big[V(\max(W-\gamma,0),A-\gamma,\tau)+(1-\kappa)\gamma-c\Big]\Bigg\}=0.\end{aligned}\tag{G.16}$$

Appendix H

Proofs for Continuous Withdrawal GMWB Variable Annuities

In this appendix we give the proofs for Lemmas 8.5, 8.6 and 8.9

H.1 Proof for Lemma 8.5

In this section we prove Lemma 8.5.

Let us define $\|V_j^n\|_\infty = \max_i |V_{i,j}^n|$. As well, let $(V_j^n)_{max} = \max_i(V_{i,j}^n)$, $(V_j^{n+})_{max} = \max_i(V_{i,j}^{n+})$, $(V_j^n)_{min} = \min_i(V_{i,j}^n)$, and $(V_j^{n+})_{min} = \min_i(V_{i,j}^{n+})$. Here we only consider the continuous withdrawal case; the discrete withdrawal case follows from the same arguments. To prove the Lemma, it is sufficient to show

$$(V_j^n)_{max} \leq \|V^0\|_\infty + A_j \ , \tag{H.1}$$

$$(V_j^{n+})_{max} \leq \|V^0\|_\infty + A_j \tag{H.2}$$

$$(V_j^n)_{min} \geq 0 \tag{H.3}$$

$$(V_j^{n+})_{min} \geq 0 \tag{H.4}$$

for all $0 \leq j \leq j_{\max}$, $0 \leq n \leq N$. We will prove inequalities (H.1-H.4) using induction.

From condition (8.3.2), it is obvious that inequalities (H.1), (H.3) hold when $n = 0$ and $0 \le j \le j_{\max}$.

Assume inequalities (H.1), (H.3) hold for $n \le n_*$ and $0 \le j \le j_{\max}$, where $n_* < N$. We next show inequalities (H.2), (H.4) hold for $n = n_*$, $0 \le j \le j_{\max}$ and then (H.1), (H.3) follow for $n = n_* + 1$, $0 \le j \le j_{\max}$.

We first consider discrete equation (8.3.3) at $n = n_*$. That is,

$$V_{i,j}^{n_*+} = \sup_{\gamma_{i,j}^{n_*} \in [0, A_j]} \left[V_{\hat{i},\hat{j}}^{n_*} + f\big(\gamma_{i,j}^{n_*}\big) \right], \quad i = 0, \ldots, i_{\max} - 1, \ j = 0, \ldots, j_{\max}. \tag{H.5}$$

According to Remark 8.3, the supremum in the right hand side of (H.5) is achieved by a control, denoted by $\bar{\gamma}_{i,j}^{n_*}$. Assume that $\max(W_i - \bar{\gamma}_{i,j}^{n_*}, 0)$ and $A_j - \bar{\gamma}_{i,j}^{n_*}$ reside within an interval $[W_l, W_{l+1}]$ and $[A_m, A_{m+1}]$, respectively, where $0 \le l < i_{\max} - 1$, $0 \le m < j_{\max}$. Then computing $V_{\hat{i},\hat{j}}^{n_*}$ using linear interpolation results in

$$V_{\hat{i},\hat{j}}^{n_*} = x_A \big[x_W V_{l,m}^{n_*} + (1 - x_W) V_{l+1,m}^{n_*} \big] + (1 - x_A) \big[x_W V_{l,m+1}^{n_*} + (1 - x_W) V_{l+1,m+1}^{n_*} \big], \tag{H.6}$$

where x_W and x_A are interpolation weights satisfying $0 \le x_W \le 1$ and $0 \le x_A \le 1$. Specifically, we have

$$x_A = \frac{A_{m+1} - (A_j - \bar{\gamma}_{i,j}^{n_*})}{A_{m+1} - A_m}. \tag{H.7}$$

Using equation (H.7) and the induction assumptions $V_{l,m}^{n_*} \le \|V^0\|_\infty + A_m$, $V_{l+1,m}^{n_*} \le \|V^0\|_\infty + A_m$, $V_{l,m+1}^{n_*} \le \|V^0\|_\infty + A_{m+1}$, $V_{l+1,m+1}^{n_*} \le \|V^0\|_\infty + A_{m+1}$, equation (H.6) leads to

$$V_{\hat{i},\hat{j}}^{n_*} \le \|V^0\|_\infty + A_j - \bar{\gamma}_{i,j}^{n_*}, \qquad \forall\ 0 \le i < i_{\max},\ 0 \le j \le j_{\max}. \tag{H.8}$$

Since $c, \kappa \ge 0$, equation (8.3.1) implies that

$$f\big(\bar{\gamma}_{i,j}^{n_*}\big) \le \bar{\gamma}_{i,j}^{n_*}, \qquad \forall\ 0 \le i < i_{\max},\ 0 \le j \le j_{\max}. \tag{H.9}$$

Equations (H.5) and (H.8-H.9) lead to (the max operator disappears since we have taken

the optimal control $\bar{\gamma}_{i,j}^{n_*}$),

$$V_{i,j}^{n_*+} = V_{\hat{i},\hat{j}}^{n_*} + f\big(\bar{\gamma}_{i,j}^{n_*}\big) \leq \|V^0\|_\infty + A_j \qquad \forall\ 0 \leq i < i_{\max},\ 0 \leq j \leq j_{\max}. \tag{H.10}$$

This proves (H.2) at $n = n_*$, $0 \leq j \leq j_{\max}$.

By the induction assumptions we have $V_{i,j}^{n_*} \geq 0$, hence from equations (H.5-H.6), we must have

$$V_{i,j}^{n_*+} \geq 0 \qquad \forall\ 0 \leq i < i_{\max},\ 0 \leq j \leq j_{\max}. \tag{H.11}$$

hence equation (H.4) holds at $n = n_*$, $0 \leq j \leq j_{\max}$.

For any $i < i_{\max}$, $0 \leq j \leq j_{\max}$, and at $n = n_* + 1$, substituting (7.3.2) into (8.3.4) gives

$$V_{i,j}^{n_*+1}\big(1 + \Delta\tau(r + \alpha_i + \beta_i)\big) - \alpha_i \Delta\tau V_{i-1,j}^{n_*+1} - \beta_i \Delta\tau V_{i+1,j}^{n_*+1} = V_{i,j}^{n_*+} \tag{H.12}$$

Let i^* be the index such that $V_{i^*,j}^{n_*+1} = (V_j^{n_*+1})_{max}$. First consider the case when $i^* < i_{\max}$. Since $r \geq 0$, and $\alpha_i \geq 0$, $\beta_i \geq 0$, as indicated by the positive coefficient condition (7.3.3), equation (H.12) implies that

$$V_{i^*,j}^{n_*+1}\big(1 + \Delta\tau(r + \alpha_{i^*} + \beta_{i^*})\big) \leq (V_j^{n_*+})_{max} + V_{i^*,j}^{n_*+1}\Delta\tau(\alpha_{i^*} + \beta_{i^*}). \tag{H.13}$$

Since we have just shown that $(V_j^{n_*+})_{max} \leq \|V^0\|_\infty + A_j$, inequality (H.13) results in

$$V_{i^*,j}^{n_*+1} \leq (V_j^{n_*+})_{max} \leq \|V^0\|_\infty + A_j. \tag{H.14}$$

Next consider the case when $i^* = i_{\max}$. Discrete equation (8.3.5) and $\|V^0\|_\infty \geq W_{\max}$ imply that

$$V_{i^*,j}^{n_*+1} = e^{-\alpha\tau^{n_*+1}} W_{\max} \leq \|V^0\|_\infty + A_j. \tag{H.15}$$

The inequality in (H.15) is due to $\alpha \geq 0$. Finally, inequalities (H.14-H.15) and the assumption $V_{i^*,j}^{n_*+1} = (V_j^{n_*+1})_{max}$ show that inequality (H.1) holds for all $0 \leq j \leq j_{\max}$

and $n = n_* + 1$. A similar argument shows equation (H.3) holds for all $0 \le j \le j_{\max}$ and $n = n_* + 1$.

H.2 Proof for Lemma 8.6

In this section, we prove Lemma 8.6.

We first consider the case when $0 < W_i < W_{i_{\max}}$, $G_r\Delta\tau < A_j \le A_{j_{\max}}$, and $0 < \tau^{n+1} \le T$. In this case, condition (8.4.15) implies that $W_i > \gamma^n_{i,j}$ for all $\gamma^n_{i,j} \in [0, G_r\Delta\tau]$. Thus, according to approximation (7.3.5), we have

$$
\begin{aligned}
&(\phi + \xi)^n_{\hat{i},\hat{j}} \\
&= \begin{cases} \phi\big(W_i - \gamma^n_{i,j}, A_j - \gamma^n_{i,j}, \tau^n\big) + \xi + O\big((\Delta W_{\max} + \Delta A_{\max})^2\big), & \gamma^n_{i,j} \in [0, G_r\Delta\tau], \\ \phi\big(\max(W_i - \gamma^n_{i,j}, 0), A_j - \gamma^n_{i,j}, \tau^n\big) + \xi + O\big((\Delta W_{\max} + \Delta A_{\max})^2\big), & \gamma^n_{i,j} \in (G_r\Delta\tau, A_j]. \end{cases}
\end{aligned}
\tag{H.16}
$$

Here we can take the term ξ out of the interpolation operation $(\cdot)^n_{\hat{i},\hat{j}}$ since it is a linear interpolation. Let us define a new variable $\hat{\gamma}^n_{i,j} = \gamma^n_{i,j}/(\Delta\tau)$. Then equation (H.16) becomes

$$
(\phi+\xi)^n_{\hat{i},\hat{j}} = \phi\big(W_i - \hat{\gamma}^n_{i,j}\Delta\tau, A_j - \hat{\gamma}^n_{i,j}\Delta\tau, \tau^n\big) + \xi + O\big((\Delta W_{\max} + \Delta A_{\max})^2\big), \ \text{ if } \hat{\gamma}^n_{i,j} \in [0, G_r].
\tag{H.17}
$$

Equation (H.17) implies

$$
\begin{aligned}
\frac{\big(\phi^{n+1}_{i,j} + \xi\big) - (\phi + \xi)^n_{\hat{i},\hat{j}}}{\Delta\tau} &= (\phi_\tau)^{n+1}_{i,j} + \hat{\gamma}^n_{i,j}(\phi_W)^n_{i,j} + \hat{\gamma}^n_{i,j}(\phi_A)^n_{i,j} \\
&\quad + O\Big(\Delta\tau + \big(\Delta W_{\max} + \Delta A_{\max}\big)^2/\Delta\tau\Big), \ \text{ if } \hat{\gamma}^n_{i,j} \in [0, G_r],
\end{aligned}
\tag{H.18}
$$

where Taylor series is used to expand $\phi\big(W_i - \hat{\gamma}^n_{i,j}\Delta\tau, A_j - \hat{\gamma}^n_{i,j}\Delta\tau, \tau^n\big)$ at (W_i, A_j, τ^n). Note that the terms in the $O(\cdot)$ expressions are bounded functions of $\hat{\gamma}^n_{i,j}$.

Assuming a discretization similar to that in Appendix B is used to discretize operator

$\mathcal{L}\phi$, then from Taylor series expansions and equation (7.3.2), we obtain that

$$\big(\mathcal{L}_h(\phi+\xi)\big)_{i,j}^{n+1} = (\mathcal{L}\phi)_{i,j}^{n+1} - r\xi + O(\Delta W_{\max}). \tag{H.19}$$

Substituting equations (H.18), (H.19), $(\phi_W)_{i,j}^n = (\phi_W)_{i,j}^{n+1} + O(\Delta\tau)$, $(\phi_A)_{i,j}^n = (\phi_A)_{i,j}^{n+1} + O(\Delta\tau)$ into $\mathcal{H}_{i,j}^{n+1}$ given in (8.4.2), then unifying the mesh size/timestep parameter for the $O(\cdot)$ terms in terms of h in (7.3.1), leads to

$$\begin{aligned}
&\mathcal{H}_{i,j}^{n+1}\big(h, \phi_{i,j}^{n+1}+\xi, \big\{\phi_{l,m}^{n+1}+\xi\big\}_{\substack{l\neq i\\ m\neq j}}, \big\{\phi_{i,j}^n+\xi\big\}\big) \\
&= \Big[\phi_\tau - \mathcal{L}\phi - \sup_{\hat\gamma_{i,j}^n\in[0,G_r]}\big[\hat\gamma_{i,j}^n - \hat\gamma_{i,j}^n\phi_W - \hat\gamma_{i,j}^n\phi_A + O(h)\big]\Big]_{i,j}^{n+1} + r\xi \\
&= \Big[\phi_\tau - \mathcal{L}\phi - \sup_{\hat\gamma_{i,j}^n\in[0,G_r]}\big(\hat\gamma_{i,j}^n - \hat\gamma_{i,j}^n\phi_W - \hat\gamma_{i,j}^n\phi_A\big)\Big]_{i,j}^{n+1} + O(h) + r\xi.
\end{aligned} \tag{H.20}$$

Here the constant for the $O(h)$ term in the first equality is a bounded function of $\hat\gamma_{i,j}^n$, that is, $O(h) = H(\hat\gamma_{i,j}^n)h$, where H is a bounded function of $\hat\gamma_{i,j}^n$. Since $\hat\gamma_{i,j}^n$ is bounded, we can move the $O(h)$ term out of the sup operator following an argument similar to Lemma D.2.

We next present an intermediate result. According to Remark 7.5, $\phi_{\hat i,\hat j}^n$ is a uniformly continuous function of $\gamma_{i,j}^n$ on $[0, A_j]$. As a result, using (7.3.1) and (7.3.5) we have

$$\begin{aligned}
&\sup_{\gamma_{i,j}^n\in(G_r\Delta\tau, A_j]}\Big[\phi_{\hat i,\hat j}^n + (1-\kappa)\gamma_{i,j}^n - c\Big] \\
&= \max_{\gamma_{i,j}^n\in[G_r\Delta\tau, A_j]}\Big[\phi_{\hat i,\hat j}^n + (1-\kappa)\gamma_{i,j}^n - c\Big] \\
&= \max_{\gamma_{i,j}^n\in[G_r\Delta\tau, A_j]}\big[\phi\big(\max(W_i-\gamma_{i,j}^n, 0), A_j-\gamma_{i,j}^n, \tau^n\big) + (1-\kappa)\gamma_{i,j}^n - c\big] + O(h^2).
\end{aligned} \tag{H.21}$$

It can be shown that $\phi\big(\max(W_i-\gamma_{i,j}^n, 0), A_j-\gamma_{i,j}^n, \tau^n\big) + (1-\kappa)\gamma_{i,j}^n - c$ is uniformly

continuous on $\gamma^n_{i,j} \in [0, A_j]$. Consequently, from (H.21) we have

$$\begin{aligned}
&\sup_{\gamma^n_{i,j}\in(G_r\Delta\tau, A_j]} \Big[\phi^n_{i,\hat{j}} + (1-\kappa)\gamma^n_{i,j} - c\Big] \\
&\quad - \sup_{\gamma^n_{i,j}\in(0, A_j]} \big[\phi\big(\max(W_i - \gamma^n_{i,j}, 0), A_j - \gamma^n_{i,j}, \tau^n\big) + (1-\kappa)\gamma^n_{i,j} - c\big] \\
&= \max_{\gamma^n_{i,j}\in[G_r\Delta\tau, A_j]} \big[\phi\big(\max(W_i - \gamma^n_{i,j}, 0), A_j - \gamma^n_{i,j}, \tau^n\big) + (1-\kappa)\gamma^n_{i,j}\big] + O(h^2) \\
&\quad - \max_{\gamma^n_{i,j}\in[0, A_j]} \big[\phi\big(\max(W_i - \gamma^n_{i,j}, 0), A_j - \gamma^n_{i,j}, \tau^n\big) + (1-\kappa)\gamma^n_{i,j}\big] \\
&= O(h).
\end{aligned} \tag{H.22}$$

Note that the subtraction of two max expressions above produces an $O(h)$ error, since the function inside the max expressions is continuous on $\gamma^n_{i,j} \in [0, A_j]$ and the difference of the optimal values of $\gamma^n_{i,j}$ for two max expressions are bounded by $G_r\Delta\tau = O(h)$. Substituting (H.19) into $\mathcal{I}^{n+1}_{i,j}$ in (8.4.3), and using (H.22) and $\phi(\max(W_i - \gamma^n_{i,j}, 0), A_j - \gamma^n_{i,j}, \tau^n) = \phi(\max(W_i - \gamma^n_{i,j}, 0), A_j - \gamma^n_{i,j}, \tau^{n+1}) + O(h)$, gives

$$\begin{aligned}
&\mathcal{I}^{n+1}_{i,j}\big(h, \phi^{n+1}_{i,j} + \xi, \big\{\phi^{n+1}_{l,m} + \xi\big\}_{\substack{l\neq i\\ m\neq j}}, \big\{\phi^n_{i,j} + \xi\big\}\big) \\
&= \phi^{n+1}_{i,j} - \sup_{\gamma^n_{i,j}\in(G_r\Delta\tau, A_j]} \Big[\phi^n_{i,\hat{j}} + (1-\kappa)\gamma^n_{i,j} - c\Big] - \kappa G_r\Delta\tau - \Delta\tau\big(\mathcal{L}\phi\big)^{n+1}_{i,j} + r\xi\Delta\tau + O(h^2) \\
&= \phi^{n+1}_{i,j} - \sup_{\gamma^n_{i,j}\in(0, A_j]} \big[\phi\big(\max(W_i - \gamma^n_{i,j}, 0), A_j - \gamma^n_{i,j}, \tau^{n+1}\big) + (1-\kappa)\gamma^n_{i,j} - c\big] + O(h),
\end{aligned} \tag{H.23}$$

where the last equality uses the fact that kG_r, $(\mathcal{L}\phi)^{n+1}_{i,j}$ and $r\xi$ are all bounded.

According to (8.4.4), (8.4.6), (H.20) and (H.23), we can write

$$\begin{aligned}
&\mathcal{G}^{n+1}_{i,j}\big(h, \phi^{n+1}_{i,j} + \xi, \big\{\phi^{n+1}_{l,m} + \xi\big\}_{\substack{l\neq i\\ m\neq j}}, \big\{\phi^n_{i,j} + \xi\big\}\big) - F_{in}\big(D^2\phi(\mathbf{x}), D\phi(\mathbf{x}), \phi(\mathbf{x}), \mathbf{x}\big) \\
&= O(h) + c(\mathbf{x})\xi, \quad \text{if } 0 < W_i < W_{i_{\max}}, G_r\Delta\tau < A_j \le A_{j_{\max}}, 0 < \tau^{n+1} \le T,
\end{aligned} \tag{H.24}$$

where $c(\mathbf{x})$ is bounded satisfying $0 \le c(\mathbf{x}) \le r$. This proves the first equation in (8.4.16).

Following similar arguments as above, we can prove the rest of equations in (8.4.16).

We omit the details here.

H.3 Proof for Lemma 8.9

In this appendix we prove Lemma 8.9.

Let us first prove (8.4.19). According to the definition of lim sup, there exist sequences $h_k, i_k, j_k, n_k, \xi_k$ such that

$$h_k \to 0, \quad \xi_k \to 0, \quad \mathbf{x}_k \equiv \big(W_{i_k}, A_{j_k}, \tau^{n_k+1}\big) \to (\hat{W}, \hat{A}, \hat{\tau}) \ \text{ as } k \to \infty, \tag{H.25}$$

and

$$\begin{aligned} &\limsup_{k\to\infty} \mathcal{G}_{i_k,j_k}^{n_k+1}\big(h_k, \phi_{i_k,j_k}^{n_k+1} + \xi_k, \big\{\phi_{l,m}^{n_k+1} + \xi_k\big\}_{\substack{l\neq i_k \\ m\neq j_k}}, \big\{\phi_{i,j}^{n_k} + \xi_k\big\}\big) \\ &= \limsup_{\substack{h\to 0,\xi\to 0, \\ \mathbf{x}\to\hat{\mathbf{x}}}} \mathcal{G}_{i,j}^{n+1}\big(h, \phi_{i,j}^{n+1} + \xi, \big\{\phi_{l,m}^{n+1} + \xi\big\}_{\substack{l\neq i \\ m\neq j}}, \big\{\phi_{i,j}^{n} + \xi\big\}\big). \end{aligned} \tag{H.26}$$

At first, we consider the case when $\hat{\mathbf{x}} \in \Omega_{in}$. Let $\Delta\tau_k$ denote the timestep corresponding to parameter h_k. Then if h_k is sufficiently small, we have $0 < W_{i_k} < W_{i_{\max}}$, $G_r\Delta\tau_k < A_{j_k} \le A_{j_{\max}}$ and $0 < \tau^{n_k+1} \le T$. According to (8.4.16), we have

$$\begin{aligned} &\mathcal{G}_{i_k,j_k}^{n_k+1}\big(h_k, \phi_{i_k,j_k}^{n_k+1} + \xi_k, \big\{\phi_{l,m}^{n_k+1} + \xi_k\big\}_{\substack{l\neq i_k \\ m\neq j_k}}, \big\{\phi_{i,j}^{n_k} + \xi_k\big\}\big) \\ &= F_{in}\big(D^2\phi(\mathbf{x}_k), D\phi(\mathbf{x}_k), \phi(\mathbf{x}_k), \mathbf{x}_k\big) + O(h_k) + c(\mathbf{x}_k)\xi_k. \end{aligned} \tag{H.27}$$

Thus, (H.26-H.27) and continuity of F_{in} (see Remark 8.8) lead to

$$\begin{aligned} &\limsup_{\substack{h\to 0,\xi\to 0, \\ \mathbf{x}\to\hat{\mathbf{x}}}} \mathcal{G}_{i,j}^{n+1}\big(h, \phi_{i,j}^{n+1} + \xi, \big\{\phi_{l,m}^{n+1} + \xi\big\}_{\substack{l\neq i \\ m\neq j}}, \big\{\phi_{i,j}^{n} + \xi\big\}\big) \\ &\le \limsup_{k\to\infty} F_{in}\big(D^2\phi(\mathbf{x}_k), D\phi(\mathbf{x}_k), \phi(\mathbf{x}_k), \mathbf{x}_k\big) + \limsup_{k\to\infty}\big[O(h_k) + c(\mathbf{x}_k)\xi_k\big] \\ &= F_{in}\big(D^2\phi(\hat{\mathbf{x}}), D\phi(\hat{\mathbf{x}}), \phi(\hat{\mathbf{x}}), \hat{\mathbf{x}}\big) \\ &= F^*\big(D^2\phi(\hat{\mathbf{x}}), D\phi(\hat{\mathbf{x}}), \phi(\hat{\mathbf{x}}), \hat{\mathbf{x}}\big), \end{aligned} \tag{H.28}$$

which verifies condition (8.4.19) for $\hat{\mathbf{x}} \in \Omega_{in}$.

We then consider the case when $\hat{\mathbf{x}} \in \Omega_{A_0} \backslash \{0\} \times \{0\} \times (0, T]$, that is, $\hat{\mathbf{x}}$ resides in region Ω_{A_0} excluding a corner line $\hat{\mathbf{x}} \in \{0\} \times \{0\} \times (0, T]$. When k is sufficiently large so that $\mathbf{x}_k$ is sufficiently close to $\hat{\mathbf{x}}$, each element in the convergent sequence $\mathbf{x}_k = (W_{i_k}, A_{j_k}, \tau^{n_k+1})$ satisfies $0 < \tau^{n_k+1} \leq T$, $0 < W_{i_k} < W_{i_{\max}}$, and either $G_r \Delta\tau_k < A_{j_k} \leq A_{j_{\max}}$, or $0 < A_{j_k} \leq G_r \Delta\tau_k$, or $A_{j_k} = 0$. Thus from (8.4.16), we have

$$\begin{aligned}
&\mathcal{G}_{i_k,j_k}^{n_k+1}\big(h_k, \phi_{i_k,j_k}^{n_k+1} + \xi_k, \big\{\phi_{l,m}^{n_k+1} + \xi_k\big\}_{\substack{l \neq i_k \\ m \neq j_k}}, \big\{\phi_{i,j}^{n_k} + \xi_k\big\}\big) \\
&= \begin{cases}
F_{in}\big(D^2\phi(\mathbf{x}_k), D\phi(\mathbf{x}_k), \phi(\mathbf{x}_k), \mathbf{x}_k\big) + O(h_k) + c(\mathbf{x}_k)\xi_k & \text{if } \ G_r\Delta\tau_k < A_{j_k} \leq A_{j_{\max}}; \\
F_{A'}\big(D^2\phi(\mathbf{x}_k), D\phi(\mathbf{x}_k), \phi(\mathbf{x}_k), \mathbf{x}_k\big) + O(h_k) + c(\mathbf{x}_k)\xi_k & \text{if } \ 0 < A_{j_k} \leq G_r\Delta\tau_k; \\
F_{A_0}\big(D^2\phi(\mathbf{x}_k), D\phi(\mathbf{x}_k), \phi(\mathbf{x}_k), \mathbf{x}_k\big) + O(h_k) + c(\mathbf{x}_k)\xi_k & \text{if } \ A_{j_k} = 0.
\end{cases}
\end{aligned} \tag{H.29}$$

From definitions of F_{A_0} and $F_{A'}$ in (8.4.8) and (8.4.13), and from $\sup_{\hat{\gamma} \in [0, A_{j_k}/\Delta\tau_k]}\big[\hat{\gamma} - \hat{\gamma}\phi_W(\mathbf{x}_k) - \hat{\gamma}\phi_A(\mathbf{x}_k)\big] \geq 0$, we observe that

$$\begin{aligned}
&F_{A'}\big(D^2\phi(\mathbf{x}_k), D\phi(\mathbf{x}_k), \phi(\mathbf{x}_k), \mathbf{x}_k\big) \\
&\leq F_{A_0}\big(D^2\phi(\mathbf{x}_k), D\phi(\mathbf{x}_k), \phi(\mathbf{x}_k), \mathbf{x}_k\big), \quad \text{if } \ 0 < A_{j_k} \leq G_r\Delta\tau_k \ .
\end{aligned} \tag{H.30}$$

As a result, (8.4.11-8.4.12) and (H.29-H.30) lead to

$$\begin{aligned}
&\limsup_{k\to\infty} \mathcal{G}_{i_k,j_k}^{n_k+1}\big(h_k, \phi_{i_k,j_k}^{n_k+1} + \xi_k, \big\{\phi_{l,m}^{n_k+1} + \xi_k\big\}_{\substack{l \neq i_k \\ m \neq j_k}}, \big\{\phi_{i,j}^{n_k} + \xi_k\big\}\big) \\
&\leq \limsup_{k\to\infty} F\big(D^2\phi(\mathbf{x}_k), D\phi(\mathbf{x}_k), \phi(\mathbf{x}_k), \mathbf{x}_k\big) + \limsup_{k\to\infty}\big[O(h_k) + c(\mathbf{x}_k)\xi_k\big] \\
&\leq F^*\big(D^2\phi(\hat{\mathbf{x}}), D\phi(\hat{\mathbf{x}}), \phi(\hat{\mathbf{x}}), \hat{\mathbf{x}}\big).
\end{aligned} \tag{H.31}$$

This together with (H.26) verify (8.4.19) for $\hat{\mathbf{x}} \in \Omega_{A_0} \backslash \{0\} \times \{0\} \times (0, T]$.

Following arguments similar to the above, we can prove (8.4.19) for the corner line $\hat{\mathbf{x}} \in \{0\} \times \{0\} \times (0, T]$ as well as for the boundary regions Ω_{W_0}, Ω_{W_m} and Ω_{τ^0}. We omit the details here.

Showing condition (8.4.20) follows in the same manner as above: we verify the condi-

tion for different regions defined in (8.4.5). Here we only show (8.4.20) for $\hat{\mathbf{x}} \in \Omega_{A_0}\backslash\{0\} \times \{0\} \times (0, T]$. Let $h_k, i_k, j_k, n_k, \xi_k$ be sequences satisfying (H.25) such that

$$\begin{aligned}
&\liminf_{k\to\infty} \mathcal{G}^{n_k+1}_{i_k,j_k}\big(h_k, \phi^{n_k+1}_{i_k,j_k} + \xi_k, \big\{\phi^{n_k+1}_{l,m} + \xi_k\big\}_{\substack{l\neq i_k \\ m\neq j_k}}, \big\{\phi^{n_k}_{i,j} + \xi_k\big\}\big) \\
&= \liminf_{\substack{h\to 0, \xi\to 0, \\ \mathbf{x}\to\hat{\mathbf{x}}}} \mathcal{G}^{n+1}_{i,j}\big(h, \phi^{n+1}_{i,j} + \xi, \big\{\phi^{n+1}_{l,m} + \xi\big\}_{\substack{l\neq i \\ m\neq j}}, \big\{\phi^{n}_{i,j} + \xi\big\}\big).
\end{aligned} \tag{H.32}$$

Then for sufficiently large k, from (8.4.16), equation (H.29) holds as discussed above. Then from definitions of F_{in} and $F_{A'}$ in (8.4.6) and (8.4.13), and from $\sup_{\hat{\gamma}\in[0,A_{j_k}/\Delta\tau_k]}\big[\hat{\gamma} - \hat{\gamma}\phi_W(\mathbf{x}_k) - \hat{\gamma}\phi_A(\mathbf{x}_k)\big] \le \sup_{\hat{\gamma}\in[0,G_r]}\big[\hat{\gamma} - \hat{\gamma}\phi_W(\mathbf{x}_k) - \hat{\gamma}\phi_A(\mathbf{x}_k)\big]$ if $0 < A_{j_k} \le G_r\Delta\tau_k$, we obtain

$$\begin{aligned}
&F_{A'}\big(D^2\phi(\mathbf{x}_k), D\phi(\mathbf{x}_k), \phi(\mathbf{x}_k), \mathbf{x}_k\big) \\
&\ge F_{in}\big(D^2\phi(\mathbf{x}_k), D\phi(\mathbf{x}_k), \phi(\mathbf{x}_k), \mathbf{x}_k\big), \quad \text{if } \ 0 < A_{j_k} \le G_r\Delta\tau_k \ .
\end{aligned} \tag{H.33}$$

Consequently, (8.4.11-8.4.12), (H.29) and (H.33) lead to

$$\begin{aligned}
&\liminf_{k\to\infty} \mathcal{G}^{n_k+1}_{i_k,j_k}\big(h_k, \phi^{n_k+1}_{i_k,j_k} + \xi_k, \big\{\phi^{n_k+1}_{l,m} + \xi_k\big\}_{\substack{l\neq i_k \\ m\neq j_k}}, \big\{\phi^{n_k}_{i,j} + \xi_k\big\}\big) \\
&\ge \liminf_{k\to\infty} F\big(D^2\phi(\mathbf{x}_k), D\phi(\mathbf{x}_k), \phi(\mathbf{x}_k), \mathbf{x}_k\big) + \liminf_{k\to\infty}\big[O(h_k) + c(\mathbf{x}_k)\xi_k\big] \\
&\ge F_*\big(D^2\phi(\hat{\mathbf{x}}), D\phi(\hat{\mathbf{x}}), \phi(\hat{\mathbf{x}}), \hat{\mathbf{x}}\big),
\end{aligned} \tag{H.34}$$

which together with (H.32) conclude (8.4.20) for $\hat{\mathbf{x}} \in \Omega_{A_0}\backslash\{0\} \times \{0\} \times (0, T]$. The other cases for (8.4.20) follow using similar steps.

Appendix I

Derivation of GMWB Variable Annuity Pricing Equation under the Discrete Withdrawal Scenario

In this appendix, we will derive the pricing equations for the annuity value under the discrete withdrawal scenario with the presence of the mutual fund management fee. In many cases, the GMWB guarantee is provided as a separate rider for an underlying investment in a mutual fund. The GMWB rider is funded by a fee which is distinct from any management fees associated with the mutual fund. The mutual fund management may be a completely separate business unit from the unit providing the GMWB guarantee, hence the mutual fund management fees are not available to fund the GMWB guarantee. Under this scenario, the unit managing the guarantee is not permitted to short the mutual fund, and uses an index proxy to hedge the guarantee [63]. Therefore, it is important to distinguish between these two sets of fees.

We will first derive the equations for the pure GMWB guarantee. We will use the same arguments used in [90] to model segregated funds. We then convert these equations into the value of the total variable annuity, which can then be related to previous work on GMWB guarantees [69, 31].

Let α_m be the proportional fee charged by the manager of the underlying mutual

fund. Let α_g be the fee used to fund the GMWB guarantee, with $\alpha_{tot} = \alpha_m + \alpha_g$. For simplicity, we will derive the equations assuming that the underlying asset follows Geometric Brownian Motion.

Consider the following scenario. The underlying asset W in the investor's account follows

$$dW = (\mu - \alpha_{tot})W dt + W\sigma dZ, \tag{I.1}$$

where μ is the drift rate and dZ is the increment of a Wiener process. We ignore withdrawals from the account in equation (I.1) for the moment. We assume that the mutual fund tracks an index $\hat{W}$ which follows the process

$$d\hat{W} = \mu\hat{W}dt + \hat{W}\sigma dZ. \tag{I.2}$$

We assume that it is not possible to short the mutual fund, so that the obvious arbitrage opportunity cannot be exploited. We further assume that it is possible to track the index $\hat{W}$ without basis risk.

Now, consider the writer of the GMWB guarantee, with no-arbitrage value $U(W, A, t)$. The terminal condition is

$$U(W, A, t = T) = \max(A(1 - \kappa) - W, 0) \tag{I.3}$$

which represents the cash flow which must be paid by the guarantee provider. The writer sets up the hedging portfolio

$$\Pi(W, \hat{W}, t) = -U(W, t) + x\hat{W}, \tag{I.4}$$

where x is the number of units of the index $\hat{W}$.

Over the time interval $t \rightarrow t + dt$, between withdrawal dates,

$$d\Pi = -\left[\left(U_t + (\mu - \alpha_{tot})WU_W + \frac{1}{2}\sigma^2 W^2 U_{WW}\right)dt + \sigma W U_W dZ\right] + x[\mu \hat{W} dt + \sigma \hat{W} dZ] + \alpha_g W dt, \tag{I.5}$$

where the term $(\alpha_g W dt)$ represents the GMWB fee collected by the hedger. Choose

$$x = \frac{W}{\hat{W}} U_W, \tag{I.6}$$

so that equation (I.5) becomes

$$d\Pi = -\left[\left(U_t - \alpha_{tot} W U_W + \frac{1}{2}\sigma^2 W^2 U_{WW}\right)dt\right] + \alpha_g W dt. \tag{I.7}$$

Setting $d\Pi = r\Pi\, dt$ (since the portfolio is now riskless) gives

$$U_\tau = \frac{1}{2}\sigma^2 W^2 U_{WW} + (r - \alpha_{tot})WU_W - rU - \alpha_g W, \tag{I.8}$$

where $\tau = T - t$. Note that equation (I.8) has the same form as that used to value segregated fund guarantees [90, 88].

At withdrawal times τ_O^k, the holder of the GMWB will maximize the value of the guarantee, so that

$$U(W, A, \tau_O^{k+}) = \sup_{\gamma^k \in [0,A]} \left[U(\max(W - \gamma^k, 0), A - \gamma^k, \tau_O^k) + f(\gamma^k) - \min(\gamma^k, W)\right]. \tag{I.9}$$

Note that

$$f(\gamma^k) - \min(\gamma^k, W) = \begin{cases} 0 & \gamma^k \leq G^k \quad ; \quad \gamma^k < W \\ -\kappa(\gamma^k - G^k) & \gamma^k > G^k \quad ; \quad \gamma^k < W \\ \gamma^k - W & \gamma^k \leq G^k \quad ; \quad \gamma^k > W \\ (\gamma^k - W) - \kappa(\gamma^k - G^k) & \gamma^k > G^k \quad ; \quad \gamma^k > W \end{cases}, \tag{I.10}$$

which represents the total cash outflows from the writer of the guarantee to the holder of the GMWB contract, i.e. cash flows required to make up any guarantee shortfall net of penalties for withdrawals above the contract amount.

Now, let $V(W, A, \tau)$ be the value of the total variable annuity contract, i.e.

$$V = U + W. \tag{I.11}$$

In other words, V is the total variable annuity value, which includes the amount in the risky account and the separate GMWB guarantee. Substituting equation (I.11) into equation (I.3) gives

$$V(W, A, \tau = 0) = \max\left(W, A(1-\kappa)\right). \tag{I.12}$$

Similarly, substituting (I.11) into (I.8) gives

$$V_\tau = \frac{1}{2}\sigma^2 W^2 V_{WW} + (r - \alpha_{tot})WV_W - rV + \alpha_m W, \tag{I.13}$$

and finally, at withdrawal times τ_O^k we obtain (from equations (I.11) and (I.9))

$$V(W, A, \tau_O^{k+}) = \sup_{\gamma^k \in [0,A]} \left[V(\max(W - \gamma^k, 0), A - \gamma^k, \tau_O^k) + f(\gamma^k) \right]. \tag{I.14}$$

Note that if $\alpha_m = 0$, then equations (I.12 - I.14) reduce to the GMWB equations (7.2.6-7.2.7) for the discrete withdrawal case.

Although at first sight the term $\alpha_m W$ on the right hand side of equation (I.13) seems counter-intuitive, we can also derive this equation assuming the following scenario. Imagine that the hedger replicates the cash flows associated with the total GMWB contract. In this case, the underlying mutual fund can be regarded as a purely virtual instrument, following process (I.1). The hedging instrument follows process (I.2), and the hedging unit pays a sales fee of $\alpha_m W$ to the mutual fund unit. In other words, rather than having the investor directly pay the proportional fees α_m to the mutual fund unit and α_g to the guarantee provider, the investor pays a proportional fee of $\alpha_{tot} = \alpha_m + \alpha_g$ to the guarantee provider, which keeps $\alpha_g W$ to hedge the guarantee and passes along $\alpha_m W$ to the mutual

fund unit. Following a similar argument as in equations (I.4-I.8), and noting that the hedger must pay the mutual fund fees $\alpha_m W$, results in equation (I.13).

Bibliography

[1] H. Ahn, A. Danilova, and G. Swindle. Storing arb. *Wilmott Magazine*, 1, 2002. 8

[2] R. Almgren and J. Lorenz. Bayesian adaptive trading with a daily cycle. *The Journal of Trading*, Fall 2006. 1

[3] A. L. Amadori. Quasi-variational inequalities with Dirichlet boundary condition related to exit time problems for impulse control. *SIAM Journal on Control and Optimization*, 43(2):570–589, 2004. 125, 138

[4] L. Andersen and J. Andreasen. Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing. *Review of Derivatives Research*, 4:231–262, 2000. 154, 156, 157

[5] R. Bansal and H. Zhou. Term structure of interest rates with regime shifts. *The Journal of Finance*, 57(5):1997–2043, 2002. 73

[6] G. Barles. Convergence of numerical schemes for degenerate parabolic equations arising in finance. In L. C. G. Rogers and D. Talay, editors, *Numerical methods in finance*, pages 1–21. Cambridge University Press, Cambridge, 1997. 2, 4, 5, 6, 36, 39, 51, 53, 55, 56, 57, 97, 105, 118, 132, 134, 136, 137, 138, 146, 201

[7] G. Barles. Nonlinear Neumann boundary conditions for quasilinear degenerate elliptic equations and applications. *Journal of Differential Equations*, 154:191–224, 1999. 50

[8] G. Barles and J. Burdeau. The Dirichlet problem for semilinear second-order degenerate elliptic equations and applications to stochastic exit time control problems. *Communications in Partial Differential Equations*, 20:129–178, 1995. 50, 51

[9] G. Barles and E. R. Jakobsen. Error bounds for monotone approximation schemes for parabolic Hamilton-Jacobi-Bellman equations. *Mathematics of Computation*, 76:1861–1893, 2007. 55

[10] G. Barles and E. Rouy. A strong comparison result for the Bellman equation arising in stochastic exit time control problems and its applications. *Communications in Partial Differential Equations*, 23:1945–2033, 1998. 50, 51

[11] G. Barles and P. E. Souganidis. Convergence of approximation schemes for fully nonlinear equations. *Asymptotic Analysis*, 4:271–283, 1991. 2, 4, 5, 6, 51, 53, 55, 56, 57, 97, 105, 118, 132, 134, 136, 137, 138, 146, 201

[12] C. Barrera-Esteve, F. Bergeret, C. Dossal, E. Gobet, A. Meziou, R. Munos, and D. Reboul-Salze. Numerical methods for the pricing of swing options: a stochastic control approach. *Methodology and Computing in Applied Probability*, 8(4):517–540, 2006. 1, 8, 9

[13] D. Bauer, A. Kling, and J. Russ. A universal pricing framework for guaranteed minimum benefits in variable annuities. Working paper, Ulm University. 106, 108, 148

[14] R. Bermejo. Analysis of a class of quasi-monotone and conservative semi-Lagrangian advection schemes. *Numerische Mathematik*, 87:597–623, 2001. 25, 31

[15] M. Bourgoing. Viscosity solutions of fully nonlinear second order parabolic equations with l^1 dependence in time and Neumann boundary conditions. Working paper, Université de Tours. 50

[16] B. Bruder and H. Pham. Impulse control problem on finite horizon with execution delay. Working paper, Universite Paris 7 Diderot, 2007. 15, 121, 203

[17] A. Budhiraja and K. Ross. Convergent numerical scheme for singular stochastic control with state constraints in a portfolio selection problem. *SIAM Journal on Control and Optimization*, 45(6):2169–2206, 2007. 1

[18] R. Carmona and M. Ludkovski. Gas storage and supply guarantees: an optimal switching approach. Working paper, Princeton University, 2005. 8, 9, 16, 58

[19] R. Carmona and M. Ludkovski. Optimal switching and application to tolling contracts. Working paper, Princeton University, 2005. 1

[20] A. Cartea and T. Williams. UK gas markets: the market price of risk and applications to multiple interruptible supply contracts. *Energy Economics*, 2007. Forthcoming. 73

[21] S. Chaumont. A strong comparison result for viscosity solutions to Hamilton-Jacobi-Bellman equations with Dirichlet condition on a non-smooth boundary and application to parabolic problems. Preprint submitted to Elsvier Science, Université Henri Poincaré I, May 2004. 50, 51

[22] Z. Chen and P. A. Forsyth. A semi-Lagrangian approach for natural gas storage valuation and optimal operation. *SIAM Journal on Scientific Computing*, 30(1):339–368, 2007. 2

[23] Z. Chen and P. A. Forsyth. A numerical scheme for the impulse control formulation for pricing variable annuities with a guaranteed minimum withdrawal benefit (GMWB). *Numerische Mathematik*, To appear. 2

[24] Z. Chen and P. A. Forsyth. Pricing hydroelectric power plants with/without operational restrictions: a stochastic control approach. In M. Ehrhardt, editor, *Nonlinear Models in Mathematical Finance*. Nova Science Publishers, To appear. 1, 9

[25] R. Cont and E. Voltchkova. A finite difference scheme for option pricing in jump diffusion and exponential Levy models. *SIAM Journal on Numerical Analysis*, 43(4):1596–1626, 2005. 55, 57

[26] R. Cont and E. Voltchkova. Integro-differential equations for option prices in exponential Levy models. *Finance and Stochastics*, 9(3):299–325, 2005. 55

[27] G. Cortazar and E. S. Schwartz. Implementing a stochastic model for oil futures prices. *Energy Economics*, 25:215–238, 2003. 73, 190, 191

[28] E. Cramer, P. Matson, and L Rubin. Common practices relating to FASB statement 133, accounting for derivative instruments and hedging activities as it relates to variable annuities with guaranteed benefits. Practice Note, American Academy of Actuaries, 2007. 148, 151, 159

[29] M. G. Crandall, H. Ishii, and P. L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bulletin of the American Mathematical Society*, 27:1–67, 1992. 35

[30] M. G. Crandall and P. L. Lions. Viscosity solutions of Hamilton-Jacobi equations. *Transactions of the American Mathematical Society*, 277:1–42, 1983. 35

[31] M. Dai, Y. K. Kwok, and J. Zong. Guaranteed minimum withdrawal benefit in variable annuities. *Mathematical Finance*, To appear. 2, 106, 107, 108, 109, 110, 111, 112, 121, 122, 123, 124, 127, 128, 138, 140, 141, 143, 145, 147, 148, 151, 161, 167, 217

[32] M. H. A. Davis and A. R. Norman. Portfolio selection with transaction costs. *Mathematics of Operations Research*, 15(4):676–713, 1990. 1, 15

[33] M. H. A. Davis, V. G. Panas, and T. Zariphopoulou. European option pricing with transaction costs. *SIAM Journal on Control and Optimization*, 31(2):470–493, 1993. 1

[34] M. H. A. Davis, V. G. Panas, and T. Zariphopoulou. European option pricing with transaction costs. *SIAM Journal on Control and Optimization*, 31(2):470–493, 1993. 15

[35] C. de Jong. The nature of power spikes: a regime-switching approach. *Studies in Nonlinear Dynamics & Econometrics*, 10(3):Article 3, 2006. 73

[36] C. de Jong and R. Huisman. Option formulas for mean-reverting power prices with spikes. Working paper, Rotterdam School of Management at Erasmus University, 2002. 73

[37] S. Deng. Pricing electricity derivatives under alternative stochastic spot price models. In *Proceedings of the 33rd Hawaii International Conference on System Sciences*, 2000. 73

[38] S. Deng and S. S. Oren. Incorporating operational characteristics and start-up costs in option-based valuation of power generation capacity. *Probability in the Engineering and Informational Sciences*, 17:151–181, 2003. 1

[39] Y. d'Halluin, P. A. Forsyth, and G. Labahn. A penalty method for American options with jump diffusion processes. *Numerische Mathematik*, 97:321–352, 2004. 66, 71

[40] Y. D'Halluin, P. A. Forsyth, and G. Labahn. A Semi-Lagrangian approach for American asian options under jump diffusion. *SIAM Journal on Scientific Computing*, 27(1):315–345, 2005. 18, 25, 53, 55, 66, 67, 71, 97

[41] Y. d'Halluin, P.A. Forsyth, and K.R. Vetzal. Robust numerical methods for contingent claims under jump diffusion processes. *IMA Journal of Numerical Analysis*, 25:65–92, 2005. 66, 71, 156

[42] J. Douglas, Jr. and T. F. Russell. Numerical methods for convection-dominated diffusion problems based on combining the method of characteristics with finite element or finite difference procedures. *SIAM Journal on Numerical Analysis*, 19:871–885, 1982. 3, 166

[43] M. Falcone and R. Ferretti. Discrete time high-order schemes for viscosity solutions of Hamilton-Jacobi-Bellman equations. *Numerische Mathematik*, 67:315–344, 1994. 23

[44] M. Falcone and R. Ferretti. Convergence analysis for a class of high-order semi-Lagrangian advection schemes. *SIAM Journal on Numerical Analysis*, 35:909–940, 1998. 25

[45] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*. Springer, 2006. 2, 15

[46] P. A. Forsyth and G. Labahn. Numerical methods for controlled Hamilton-Jacobi-Bellman PDEs in finance. *Journal of Computational Finance*, 11(2):1–44, Winter 2007/2008. 2, 4, 9, 27, 53, 55, 97, 167, 183, 195

[47] P. A. Forsyth and K. R. Vetzal. Quadratic convergence for valuing american options using a penalty method. *SIAM Journal on Scientific Computing*, 23:2096–2123, 2000. 121, 167

[48] N. C. Framstad, B. Oksendal, and A. Sulem. Optimal consumption and portfolio in a jump diffusion market with proportional transaction costs. *Journal of Mathematical Economics*, 35:233–257, 2001. 15

[49] M. B. Giles and R. Carter. Convergence analysis of Crank-Nicolson and Rannacher time-marching. *Journal of Computational Finance*, 9(4):89–112, 2006. 59

[50] S. Gray. Modeling the conditional distribution of interest rates as a regime switching process. *Journal of Financial Economics*, 42:27–62, 1996. 73

[51] J. Hamilton. Analysis of time series subject to changes in regime. *Journal of Econometrics*, 45:39–79, 1990. 73

[52] M. R. Hardy. A regime-switching model of long-term stock returns. *North American Actuarial Journal*, 5(2):41–53, 2001. 73

[53] C. He, J.S. Kennedy, T. Coleman, P.A. Forsyth, Y. Li, and K. Vetzal. Calibration and hedging under jump diffusion. *Review of Derivatives Research*, 9:1–35, 2006. 157

[54] T. Ho, S. Bin Lee, and Y.S. Choi. Practical considerations in managing variable annuities. working paper, Thomas Ho Company, 2005. 149, 160

[55] R. Huisman and R. Mahieu. Regime jumps in electricity prices. *Energy Economics*, 25:425–434, 2003. 73

[56] K. Ishii. Viscosity solutions of nonlinear second order elliptic PDEs associated with impulse control problems II. *Funkcialaj Ekvacioj*, 38:297–328, 1995. 138

[57] P. Jaillet, E. I. Ronn, and S. Tompaidis. Valuation of commodity-based swing options. *Management Science*, 50(7):909–921, July 2004. 72, 73, 79

[58] J.S. Kennedy, P.A. Forsyth, and K. Vetzal. Dynamic hedging under jump diffusion with transaction costs. To appear in Operations Research. 157

[59] R. Korn. Portfolio optimization with strictly positive transaction costs and impulse control. *Finance and Stochastics*, 2:85–114, 1998. 1, 121

[60] R. Korn. Some applications of impulse control in mathematical finance. *Mathematical Methods of Operations Research*, 50:493–518, 1999. 121, 125, 202, 203

[61] H. J. Kushner and P. Dupuis. *Numerical methods for stochastic control problems in continuous time*. Springer-Verlag, 2001. 1

[62] T. L. Lai and T. W. Lim. A new approach to singular stochastic control in optimal investment and hedging in the presence of transaction costs. Working paper, Stanford University, 2006. 1

[63] M. Le Roux. Private communication. Director of product development, ING Institutional Markets, 2000. 148, 158, 217

[64] H. Li. Adaptive wavelet collocation methods for option pricing PDEs. PhD Thesis, University of Calgary, August 2006. 2, 8, 9

[65] P. L. Lions. Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations. *Communications in Partial Differential Equations*, 8:Part I, 1101–1134, Part II, 1229–1276, 1983. 35

[66] M. Ludkovski. Financial hedging of operational risk. Working paper, University of Michigan, 2006. 1

[67] M. Manoliu and S. Tompaidis. Energy futures prices: term structure models with Kalman filter estimation. *Applied Mathematical Finance*, 9(1):21–43, 2002. 73

[68] C. Martini and C. Patry. Variance optimal hedging in the Black-Scholes model for a given number of transactions. INRIA Technical Report 3767, 1999. 1

[69] M. A. Milevsky and T. S. Salisbury. Financial valuation of guaranteed minimum withdrawal benefits. *Insurance: Mathematics and Economics*, 38:21–38, 2006. 107, 108, 121, 145, 148, 151, 161, 164, 217

[70] B. Øksendal and A. Sulem. Optimal consumption and portfolio with both fixed and proportional transaction costs. *SIAM Journal on Control and Optimization*, 40(6):1765–1790, 2002. 1, 15, 121, 125, 138

[71] M. Pemy and Q. Zhang. Optimal stock liquidation in a regime switching model with finite time horizon. *Journal of Mathematical Analysis and Applications*, 321:537–552, 2006. 5, 95, 97, 105

[72] H. Pham. On some recent aspects of stochastic control and their applications. *Probability Surveys*, 2:506–549, 2005. 1, 2, 15, 39, 121, 125

[73] D. Pilopović. *Energy Risk*. McGraw-Hill, New York, 1998. 8, 13, 75

[74] O. Pironneau. On the transport diffusion algorithm and its applications to the Navier-Stokes equations. *Numerische Mathematik*, 38:309–332, 1982. 3, 166

[75] D. M. Pooley, P. A. Forsyth, and K. R. Vetzal. Numerical convergence properties of option pricing PDEs with uncertain volatility. *IMA Journal of Numerical Analysis*, 23:241–267, 2003. 2, 51, 59

[76] R. Rannacher. Finite element solution of diffusion problems with irregular data. *Numerische Mathematik*, 43:309–327, 1984. 59, 61

[77] G. Schindlmayr. A regime-switching model for electricity spot prices. Working paper, EnBW Trading GmbH, 2005. 73

[78] E. S. Schwartz. The stochastic behavior of commodity prices: implications for valuation and hedging. *The Journal of Finance*, 52(3):923–973, 1997. 72, 73, 76, 78, 190

[79] I. B. Tahar, H. M. Soner, and N. Touzi. Modelling continuous-time financial markets with capital gains taxes. Working paper, TU Berlin and CRESET, Paris, 2005. 1

[80] M. Thompson, M. Davison, and H. Rasmussen. Natural gas storage valuation and optimization: a real options application. Working paper, University of Western Ontario, July 2003. 2, 8, 9, 10, 11, 16, 28, 31, 60, 61, 62

[81] M. Thompson, M. Davison, and H. Rasmussen. Valuation and optimal operation of electric power plants in competitive markets. *Operations Research*, 52(4):546–562, 2004. 1

[82] C. Tseng and G. Barz. Short-term generation asset valuation: a real options approach. *Operations Research*, 50(2):297–310, 2002. 1

[83] V. L. Vath, M. Mnif, and H. Pham. A model of optimal portfolio selection under liquidity risk and price impact. *Finance and Stochastics*, 11:51–90, 2007. 1, 15, 121, 125, 138

[84] T. Ware. Swing options in a mean-reverting world. Presentation at: Stochastic Calculus and its applications to Quantitative Finance and Electrical Engineering, a conference in honour of Robert Elliott, Calgary, July 2005. 8, 9, 13, 15

[85] T. Ware and H. Li. Swing options with continuous exercise. Presentation at the Canadian Mathematical Society meeting, Calgary, June 2006. 8, 9, 13, 15

[86] P. Wilmott. *Derivatives: the theory and practice of financial engineering*. John Wiley & Sons, West Sussex, England, 1998. 177

[87] H. Windcliff, P. A. Forsyth, and K.R. Vetzal. Analysis of the stability of the linear boundary condition for the Black-Scholes equation. *Journal of Computational Finance*, 8(1):65–92, 2004. 91

[88] H. Windcliff, P.A. Forsyth, M.K. Le Roux, and K.R. Vetzal. Understanding the behaviour and hedging of segregated funds offering the reset feature. *North American Actuarial Journal*, 6:107–125, 2002. 148, 219

[89] H. Windcliff, J. Wang, P. A. Forsyth, and K. R. Vetzal. Hedging with a correlated asset: solution of a nonlinear pricing pde. *Journal of Computational and Applied Mathematics*, 200:86–115, 2007. 36

[90] H.A. Windcliff, P.A. Forsyth, and K.R. Vetzal. Valuation of segregated funds: shout options with maturity extensions. *Insurance: Mathematics and Economics*, 29:1–21, 2001. 148, 217, 219

[91] Z. Xu. Stochastic models for gas prices. Master's thesis, University of Calgary, 2004. 73, 75, 76, 78, 190

[92] V. I. Zakamouline. A unified approach to portfolio optimization with linear transaction costs. *Mathematical Methods of Operations Research*, 62:319–343, 2005. 124

[93] R. Zvan, P. A. Forsyth, and K. R. Vetzal. Discrete Asian barrier options. *Journal of Computational Finance*, 3:41–68, 1999. 31