Optimal Switching with Applications to Energy Tolling Agreements

Michael Ludkovski

A Dissertation

Presented to the Faculty

of Princeton University

in Candidacy for the Degree

of Doctor of Philosophy

Recommended for Acceptance

by the Department of

Operations Research and Financial Engineering

June 2005

Abstract

We consider the problem of optimal switching with finite horizon. This special case of stochastic impulse control naturally arises during analysis of operational flexibility of exotic energy derivatives. The current practice for such problems relies on Markov decision processes that have poor dimension-scaling properties, or on strips of spark spread options that ignore the operational constraints of the asset.

To overcome both of these limitations, we propose a new framework based on recursive optimal stopping. Our model demonstrates that the optimal dispatch policies can be described with the aid of 'switching boundaries', similar to standard American options. In turn, this provides new insight regarding the qualitative properties of the value function.

Our main contribution is a new method of numerical solution based on Monte Carlo regressions. The scheme uses dynamic programming to simultaneously approximate the optimal switching times along all the simulated paths. Convergence analysis is carried out and numerical results are illustrated with a variety of concrete examples. We then benchmark and compare our scheme to alternative numerical methods. On a mathematical level, we contribute to the numerical analysis of reflected backward stochastic differential equations and quasi-variational inequalities. The final part of the dissertation proposes fruitful extensions to tackle other financial problems such as gas storage, exhaustible resources, hedging supply guarantees and energy risk management.

Acknowledgements

I would like to begin by thanking René Carmona who introduced me to the universe of energy derivatives. This dissertation would not have existed without the numerous late night meetings we have had over the past three years. I am indebted to Erhan Çinlar and Ronnie Sircar who together with René formed my Ph.D. committee and gave much valuable advice over the years. They have also taught me most of the probability theory I know and deeply influenced my view of financial mathematics. I would like to specially thank Savas Dayanik for agreeing to be my dissertation reader and carefully checking this manuscript.

I am grateful to Monique Jeanblanc, Bernard Lapeyre and Nizar Touzi for useful and stimulating discussions that have been key in guiding my research on tolling agreements. Thanks also go out to Valdo Durrleman, Masa Egami, Jamey Thompson and all my other colleagues at Bendheim and ORFE who made my graduate experience at Princeton so much fun. Finally, I would like to thank my parents and entire family for their support; this thesis is dedicated to them.

To My Parents.

Contents

List of Figures

List of Tables

Notation Index

Chapter 1

Introduction

This dissertation is concerned with finding optimal policies for exercise of operational flexibility over energy assets. This is one of the fundamental problems faced by participants in the emerging energy markets. To reduce the large liquidity risk present in these markets trading firms are increasingly seeking control of generating assets like power plants or gas storage facilities. The ownership is transferred by signing temporary lease agreements and it is of interest to efficiently value and hedge such *tolling contracts* given the volatile gas and electricity prices.

In this work to solve the outlined problem we adopt a framework of stochastic control in continuous time. The commodity prices are modelled as stochastic processes and the owner exercises her managerial options by controlling the production mode of the asset. This approach was pioneered by Brennan and Schwartz [13] over twenty years ago, but only recently received attention. However, the dynamic setting is the only way to fully capture the interplay between flexibility and uncertainty. The method is technical, but the intuitiveness is not lost. We obtain explicit results which have simple interpretations and lend themselves well to practical implementations.

Our key insight is that optimal switching is a special case of stochastic impulse control that is efficiently reduced to a series of *recursive optimal stopping* problems.

At any given instant in time, the optimal policy is determined by the simple decision between 'continue in the same production regime' and 'switch to the best alternative'. From this idea, we develop a new numerical algorithm based on Monte Carlo simulations and Dynamic Programming. Our main motivation is robust implementation and applicability to a wide range of financial engineering settings.

Besides a numerical scheme, the second contribution of this work is a new qualitative analysis of optimal switching. Under conditions to be specified later on, we prove that the optimal switching decisions can be fully described with the aid of switching boundaries. These boundaries are smooth and divide the state space into connected components. Our analysis is entirely probabilistic, avoiding the difficulties of the traditional quasi-variational framework. In particular, we do not make any ad hoc guesses about the optimal policy. Instead we rely on the theory of reflected backward stochastic differential equations and the new technique of maturity randomization. Maturity randomization studies finite horizon problems as limits of an iterative sequence of infinite horizon ones and naturally fits in our framework.

Compared to existing literature, we consider more general stochastic price processes and more importantly look at finite horizon problems. Rather than valuing a project on an infinite time interval, we concentrate on contracts with fixed expiry dates. This is closer to reality but makes the problem much more difficult. Time has to explicitly enter into all the calculations and time decay may become significant.

To the best of our knowledge this work is the first to apply continuous time impulse control in the context of tolling agreements for energy assets. Accordingly, we highlight the differences and advantages of our model versus existing methods in a variety of numerical and qualitative examples. We especially concentrate on the strip of options approximation that is widely used nowadays by practitioners.

The organization of this thesis is as follows. Chapter 2 begins by introducing the financial motivation for our problem and carefully formulating the precise mathe-

matical framework. We then proceed to Chapter 3 which presents the main concept of converting optimal switching problem into recursive optimal stopping. Building on the basic Theorems 1 and 2 we then discuss the analytical properties of the value function. The chapter ends by recalling the alternative approaches based on quasi-variational inequalities and backward stochastic differential equations and the additional insights they provide.

Chapter 4 is the backbone of this dissertation and describes the numerical approach we use to solve the problem. The main regression-based dynamic programming algorithm is presented in Section 4.1. In Section 4.3 we then systematically discuss convergence issues with separate analysis of each source of error. Moving on to practical issues, Section 4.4 provides several representative examples and Section 4.5 compares our algorithm to alternative numerical methods. In a different vein, Section 4.7 discusses the differences of our framework versus approaches currently in use by practitioners.

The final Chapter 5 is devoted to extending our framework to additional cases of interest. We begin by discussing a model of gas storage in Section 5.1. While similar to dispatch of a power plant, this problem has past-dependent state variables that complicate implementation. We propose two possible solutions in Section 5.1.1 and compare their performance on a couple more examples. Gas storage naturally leads to other generalizations. Specifically, we discuss supply guarantees in Section 5.2, exhaustible resources in Section 5.3, and nonlinear risk preferences in Section 5.5. The last extension in turn leads to a new proposal in Section 5.6 for integration of intrinsic risk preferences and vanilla instruments with managerial flexibility. Combined we obtain a risk management framework that brings us full circle and allows us to resolve both pricing and hedging of tolling agreements. The concluding Section 5.7 summarizes our work and discusses avenues for further research.

Chapter 2

Problem Setup

2.1 Tolling Agreements

The energy industry is fundamentally very capital intensive with project costs of seven or eight digits. A typical fossil fuel power plant costs in the hundreds of millions of dollars and may take three to five years to build. Oil refineries or gas storage facilities also require enormous capital outlays. As a result, the physical stock is owned by a few very large firms whose expertise is in building and maintaining the infrastructure. From a financial point of view however, energy assets are really a play on the *spread* between two different commodities. To put it another way, energy assets are simply black boxes that convert a fuel commodity into another fuel commodity. Thus, a power plant converts gas or oil into electricity, a refinery converts crude oil into gasoline and jet fuel, and a storage facility converts gas today into gas six months from now.

Because energy markets are not sufficiently liquid and efficient, access to the physical 'equipment' has extra benefits that cannot be attained otherwise. In particular, many commodity contracts require physical settlement which necessitates actual ownership of an asset. Consequently, many energy trading firms have a vested

interest in owning energy assets. To circumvent the capital intensive side of the business, the idea of a *tolling agreement* was invented.

For a trader, a tolling agreement is a call on power with a gas-linked strike price. Thus, if the gas/power spread is large enough, the trader will run the plant turning a profit. In general, the buyer has the right to plant output at his discretion, subject to pre-specified exercise rules. Since the latter can be quite complex, for our purposes a tolling agreement is simply any temporary contract between the permanent owner of an asset and another agent that allows that agent to claim ownership and management of the output. In short, the agent 'rents' the asset from the owner. The arrangement permits the owner to concentrate on maintenance and development while allowing the agent to hedge and/or speculate in the gas and power derivatives markets.

Tolling agreements come in a variety of shapes and flavors. Being a structured deal, each contract is different and uniquely tailored to the participants' needs. There even exist reverse tolls which allow power producers to virtually convert their electricity back into gas. In this thesis we abstract from the specifics and concentrate on the optimal behavior of the renter in a general mathematical framework. For concreteness, from now on we discuss the case of a tolling agreement for a gas-fired power plant in a de-regulated market. The agent is exposed to fluctuating fuel and electricity prices and would like to derive the maximum value from the plant. This is achieved by optimizing the dispatching policy, i.e. deciding when the plant is running and when it is offline. These decisions are made dynamically, as time and market conditions evolve. We shall assume that the market is infinitely liquid and the agent is a price taker, so that her actions do not have any effect on the prices. The last assumption is very strong, however very few financial models can take into account price impact in a succinct manner.

2.2 Operating Strategies

Consider an operator in charge of a peaking combined cycle gas turbine (CCGT) power plant. This is a medium size new technology power plant characterized by short ramping times. Hence it only takes $2-4$ hours to bring the plant online, in contrast to baseload units that require days of powering up. Several hundred of such units have been built in North America with typical capacity of $40-200MWh$. As the name suggests, a CCGT plant is made up of several gas-fired turbines that allow for variable levels of output.

If the operator wishes to run the plant, she buys natural gas, converts it into electricity and sells the output on the market. The conversion ratio is called the *heat rate*. More precisely, the heat rate $\overline{HR}$ specifies the number of millions of British thermal units (MMBtu) of gas needed to produce one megawatt-hour MWh.[1] Thus, the higher the heat rate the less efficient the conversion. Typical heat rates are about $9-12MWh/MMBtu$. To formulate the financial setting, let (P_t) be the price process of electricity, (G_t) be the price process of gas, K the operating costs and Cap the capacity of the plant. We remain vague about the precise meaning of (P_t) and (G_t). They could be spot prices, but they could also very well be day-ahead prices if all the commitments are made on a 24-hour basis. The revenue rate from running the plant is then given by the *spark spread*, $Cap\cdot(P_t-\overline{HR}\cdot G_t-K)\cdot dt$. In other words the spark spread pays the difference between the market price of power and the market price of gas needed to produce this power. The remaining margin, which may be negative, is the current economic rent of owning the plant.

Suppose that besides running the plant at full capacity with heat rate $\overline{HR}$ or keeping it completely off-line, there also exist a total of $M-1$ intermediate operating modes or regimes, corresponding to different subsets of turbines running. In

[1] In Europe, heat rates are quoted in megawatt-hours per giga-joule GJ. We use the American units.

principle, the plant may have a continuous spectrum of operating regimes. However, some specific output levels are likely to be more operationally stable so that our model is an acceptable simplification. To each mode we associate the corresponding marginal heat rate HR_i, such that they are ranked in increasing order of inefficiency, $0 = HR_0 \leqslant HR_1 \leqslant HR_2 \leqslant \ldots \leqslant HR_M$, $\sum_i HR_i = \overline{HR}$. The marginal benefit is always decreasing, causing dis-economies of scale. One reason this happens is due to increasing losses from heat dissipation. Each mode also has its own rate of O&M costs K_i. The rate of payoff in regime m is then given by

$$\psi_m(P_t, G_t) \triangleq Cap\Big(\frac{m}{M} \cdot P_t - \sum_{i=0}^{m} HR_i \cdot G_t - K_i\Big). \tag{2.1}$$

In general, we denote by $X_t = (P_t, G_t)$ the stochastic $\mathbb{R}^2$-valued driving process and by $\psi_m(t, X_t)$ the respective payoff rates. Subject to integrability conditions to be specified later on, ψ_m is allowed to be any time dependent Lipschitz-continuous functional on the domain of (X_t). In particular, ψ_m might incorporate time discounting at some rate r. However, to avoid clutter from now on we will not write out explicitly either e^{-rt} or the K_i's.

Changing an output level is costly. It often requires extra fuel and various overhead costs. Moreover, decisions must be synchronized to ensure operational stability. For ease of presentation, let us assume that the startup and shutdown costs are equal to C for each unit, so that the cost of switching from mode i to mode j is $C|i - j|$. At this point we ignore the time delay effect of having to gradually 'ramp-up' and 'ramp-down' the turbine, an issue that we revisit in Section 4.1.1. In the general setting we will also write $C_{i,j}$ and we can even allow dependence on time and current state $C_{i,j}(t, X_t)$. The switching costs are discrete with $C_{i,j} > \epsilon > 0$, for all i, j and satisfy the triangle inequality $C_{i,j} \leqslant C_{i,k} + C_{k,j}$ for any intermediate regime k. We further assume that the cost of not running a plant is zero. Fixed expenses are ignored because those are deterministic sunk costs representing an additional component of being an operator and must be paid regardless of dispatching policy.

The problem we shall investigate consists of optimal use of the power plant on a finite time horizon $[0,T]$. We have in mind a relatively short time period, such as six months or at most a year. The optionality of running the plant comes *only* from startup/shutdown decisions. Thus, the operating strategies in our setting consist of the double sequences $u = (\xi, \mathcal{T})$ where ξ_k taking values in $\mathbb{Z}_M \triangleq \{0, \cdots, M-1\}$ are the successive modes chosen by the strategy u and $0 \leqslant \tau_{k-1} \leqslant \tau_k \leqslant T$ are the switching times. Since the only observable is (X_t) itself, we require τ_k to be $\mathcal{F}^X$-stopping times, where $\mathcal{F}^X_t = \sigma(X_s \colon 0 \leqslant s \leqslant t)$ is the filtration generated by (X_t). In general, several successive switches are allowed so that $\tau_k = \tau_{k+1}$ is possible. However, due to the assumption made in the previous paragraph about subadditive costs, multiple instantaneous switches are suboptimal. Note that this assumption is without loss of generality because one can simply re-define $C_{i,j} = \min_k(C_{i,k} + C_{k,j})$ without changing the structure of the problem. Consequently, one can think of u as an $\mathcal{F}^X$-adapted piecewise-constant and càdlàg (right-continuous with left-limits) process where u_s denotes the operating mode at time s. The jumps of u are precisely described by τ_k. The total reward up to fixed final time T for such control u is

$$H(x,i,[0,T];u)(\omega) \triangleq \int_0^T \psi_{u_s}(s, X_s)\, ds - C \int_0^T |du|_s, \qquad X_0 = x, u_0 = i. \tag{2.2}$$

The second term above is a convenient shorthand notation for counting the cumulative switching costs corresponding to u. In full generality it should be written as $\sum_{\tau_k < T} C(u_{\tau_k -}, u_{\tau_k}; \tau_k, X_{\tau_k -})$ which we find too cumbersome. Observe that the reward functional $H(x,i,[0,T];\cdot)$ is time additive which will be crucial for structural properties of the problem.

2.3 Control Problem

Let $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}^X_t), \mathbb{P})$ be a stochastic basis. As usual in finance, the measure $\mathbb{P}$ corresponds to the risk-neutral valuation measure. Let $\mathcal{U}$ (respectively $\mathcal{U}(t)$) be the set of all *acceptable* controls on $[0,T]$ (resp. on $[t,T]$). It consists of all adapted

càdlàg $\mathbb{Z}_M$-valued processes u of a.s. finite variation on $[t,T]$. The last condition means that we require the number of switches to be finite almost surely. Thus, there exists a random integer $N_T = N_T(u)$ such that $\tau_{N_T} = T$. Alternatively, $\mathbb{P}[\tau_n < T \ \forall n \geqslant 0] = 0$. This restriction is superfluous if the expected maximum gain is finite a.s., $\mathbb{P}^x[\int_0^T \max_i \psi_i(s, X_s)\, ds = +\infty] = 0$. Indeed, for any strategy u, on the set $\mathbf{A}^u_\infty \triangleq \{N_T(u) = \infty\}$ the switching costs are infinite, and hence $H(x, i, [0,T]; u) = -\infty$ on $\mathbf{A}^u_\infty$. As a result if $\mathbb{P}^x(\mathbf{A}^u_\infty) > 0$ then the mean expected loss from strategy u is infinite and the latter is clearly not optimal. We conclude that only strategies u for which $\mathbb{P}^x(\mathbf{A}^u_\infty) = 0$ need to be considered. From an economic perspective the assumption also makes a lot of sense— switching a plant a thousand times a day is impossible.

The optimal switching problem we will investigate may now be rigorously written as finding

$$J(t,x,i) = \sup_{u \in \mathcal{U}(t)} J(t,x,i;u), \tag{2.3}$$

$$\text{where} \qquad J(t,x,i;u) \triangleq \mathbb{E}\big[H(x,i,[t,T];u) \big| \, X_t = x, u_t = i\big].$$

In full,

$$J(t,x,i) = \sup_{u \in \mathcal{U}(t)} \mathbb{E}\left[\int_t^T \psi_{u_s}(s, X_s)\, ds - \sum_{t \leqslant \tau_k < T} C(u_{\tau_k -}, u_{\tau_k}) \Big| \, X_t = x, u_t = i\right]. \tag{2.4}$$

In Section 3.3 we will impose further conditions on the ingredients and verify that the above control problem is sensible. In words, $J(t,x,i)$, which is called the value function, is the conditional maximum expected value for running the plant on $[t,T]$ given the initial value of $X_t = x$ and the initial regime $u_t = i$. In economic terms, $J(t,x,i)$ represents the net present value of all future profit flows given optimal behavior henceforth. Our goal is to numerically compute the value function $J(t,x,i)$, describe its qualitative properties and characterize the optimal switching policy $u^* = (\xi^*, \mathcal{T}^*)$, if one exists, that achieves the supremum in (2.4). The last item is crucial

for practical applications where the agent needs an easily implementable dispatch strategy.

Mean Reversion and the Spark Spread

In contrast to most existing finance literature, our main case of interest is a mean-reverting driving process (X_t). Indeed, unless (X_t) is mean-reverting and recurrent, the problem will become degenerate as the drift would eventually make one regime preferable to all the rest. Furthermore, for our problem, empirical data [27, 38] strongly suggest that the spark spread $P_t - \overline{HR} \cdot G_t$ is stationary and fluctuates around its long-term mean corresponding to the average flow of economic rent from a power plant.

An abbreviated alternative that will be used for some of our illustrations is to assume that (X_t) is one-dimensional. This corresponds to direct modeling of the spark spread, see for example [15]. In that case we will assume that the payoff rates are of the form $\psi_m(X_t) = m \cdot X_t - \sum_{i=0}^{m} \beta_i \cdot K_g$ where $\beta_0 = 0, \beta_1 \geqslant \ldots \geqslant \beta_M$. This is the one-dimensional analogue of decreasing efficiency of production.

2.4 Relation to Existing Literature

Optimal switching is a rather old problem that has been studied by both economists and mathematicians. Mathematically, it is a special case of impulse control and the latter has been extensively analyzed since late 1970s. We review relevant results in this direction in Section 3.6. Economically, several other financial applications have been looked at. The largest body of related literature treats problems of partially reversible investment encountered in real options. In this setting the agent is a firm facing several investment projects with uncertain dynamic value (X_t) that it can start and suspend. The control is therefore composed of investment times τ_k and choice/size of project to start ξ_k. For example, such optimal capacity management

by an industrial firm has been studied by A. Øksendal [60] and extended to the case of partial disinvestment by Guo and Pham [43].

Another related problem is optimal dividend payout by a corporation. Given stochastic firm value (X_t), the objective is to find the best method of distributing the wealth to shareholders. Thus, the controls are dividend times τ_k and dividend amounts ξ_k. Mathematical treatment of this problem first appeared in Jeanblanc and Shiryaev [48]. Extensions including non-zero recovery at default and mean-reverting (X_t) have been recently studied by Boguslavsky [9] and Cadenillas et al. [14] respectively.

Directly related to our setting we should foremost mention the recent work by Hamadène and Jeanblanc [44]. Their probabilistic approach has been the inspiration for this research. However, in their model there are only two operating regimes and they spend little time discussing numerical implementation. The traditional variational approach to optimal switching originated with Brekke and Øksendal [12] who considered a geometric Brownian motion for (X_t) and infinite horizon. Another similar work is by Yushkevich [71]. He considers the switching problem in discrete time for a general recurrent Markov chain with countable state space. Again, there are only two regimes and no running rewards which allows for geometric characterization of the value function.

The original paper on using stochastic control for commodity asset dispatch is due to Brennan and Schwartz [13]. They used the methodology to price a copper mine, once more with geometric Brownian motion for commodity price (X_t). Further extensions such as abandonment and initial start decisions have appeared in a series of recent papers by Zervos and various co-authors [28, 57, 72].

From an economic perspective, the thrust has been to show that stochasticity of the (X_t) state process together with positive switching costs cause investment delay and the appearance of the hysteresis band [26]. This means that the owner

will forgo possible small gains (or respectively suffer small losses) due to the large outlay required to make a switch. Thus, an operator may continue to run a plant at a loss if he has enough hope that the prices will soon increase. Similarly, once prices are on the rise, the manager will be reluctant to bring the plant online immediately, delaying his decision. Overall, the fixed switching costs in the face of uncertain revenue streams induce risk-aversion and reduce net present value (NPV). There is also a competing effect, termed the timing option [1, 28]. Because the manager is able to time his decisions he can minimize his losses in unfavorable conditions, as well maximize them in good ones. As a result he derives extra benefits that essentially correspond to the time premium for American option rights and are often a crucial ingredient of the project NPV.

With the exception of Hamadène and Jeanblanc [44], all other 'mathematical' papers mentioned so far concentrate on explicit solutions and restrict their attention to infinite horizon and one-dimensional time-homogeneous diffusions for (X_t). By abstracting to these stylized setting, it is possible to find analytical solutions which are used to demonstrate key features. However, the shortcoming is that most models are impractical, and it is not clear how to implement them in a real-life situation. One of the motivations behind this thesis was to redress this problem and focus on numerical robustness and scalability.

Chapter 3

Recursive Optimal Stopping

3.1 Setup and Assumptions

We begin by stating the technical setup we will work in. For our driving process we take (X_t) to be an Itô diffusion[1] on an open subset $E \subseteq \mathbb{R}^d$. We do not specify here boundary conditions and our typical space is in fact the entire $\mathbb{R}^d$. In that case the dynamics of (X_t) are representable by a stochastic differential equation (SDE)

$$dX_t = \mu(X_t)\,dt + \sigma(X_t) \cdot dW_t, \tag{3.1}$$

where W_t is a standard Wiener process on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. The filtration $\mathbb{F}$ is Brownian and satisfies the *usual* conditions. We assume that $\mathcal{F}_0$ is trivial and our initial laws are always deterministic point-masses at some x giving rise to conditional probability laws $\mathbb{P}^x$. We will write $(X_s^{t,x})$ to indicate the process conditional on $X_t = x$. Moreover, we make a standing assumption that the transition law $P_t(x, y)$ of (X_t) is non-degenerate in the entire domain: $\int_0^\infty P_t(x, y)\,dt > 0,\ \forall x, y \in E$. As usual, E is equipped with its Borel σ-algebra $\mathcal{B}(E)$ and regularity implies that the former coincides with the intrinsic topology of (X_t). For vector norm we use the Euclidean $\|x\| = \sum_i x_i^2$.

[1] Generalizations are possible, see Section 5.4.

Throughout we assume that the SDE (3.1) is non-degenerate, i.e. the eigenvalues of $\sigma(x)$ are bounded away from zero, $\inf_{x\in E}|\lambda_{\min}(\sigma(x))|>0$, and has a unique strong solution. One sufficient condition is for μ and σ to be locally Lipschitz:

$$\|\mu(x)-\mu(y)\|+\|\sigma(x)-\sigma(y)\|<K_N\|x-y\|,\qquad \forall\|x\|,\|y\|\leqslant N. \tag{3.2}$$

Both μ and σ can be time dependent and we suppress dependence on t purely for convenience. Let

$$\mathscr{S}_T^p\triangleq\big\{Z\colon Z_t\in\mathcal{F}_t,\ \mathbb{E}\sup_{t\in[0,T]}|Z_t|^p<\infty\big\},\qquad p\geqslant 1. \tag{3.3}$$

We then make

Assumption 1. *For all m, the reward function $\psi_m\colon[0,T]\times E\to\mathbb{R}$ is Borel, continuous and locally Lipschitz in (t,x). Furthermore, $\psi_m(\cdot,X_\cdot)\in\mathscr{S}_T^2$.*

The last condition is satisfied, for instance, if $(X_t)\in\mathscr{S}_T^1$ and all the rewards are of quadratic growth, $|\psi_m(t,x)|<C(1+\|x\|^2)$.

Our canonical example for (X_t) is a d-dimensional exponential Ornstein-Uhlenbeck process, namely

$$\frac{dX_t^i}{X_t^i}=\kappa^i(\theta^i-\log X_t^i)\,dt+\Sigma\cdot dW_t,\qquad i=1,\dots,d, \tag{3.4}$$

$$\text{or}\qquad d(\log X_t^i)=\kappa^i(\theta^i-\frac{\sigma_i^2}{2\kappa^i}-\log X_t^i)\,dt+\Sigma\cdot dW_t,\qquad X_0^i=x^i,$$

where W_t is a d-dimensional Brownian motion and $\Sigma\in\mathbb{R}^{d\times d}$ is a constant non-degenerate volatility matrix. Thus, the dependence between the components of (X_t) is only through the correlation in the driving Wiener process. The attractiveness of this model is that $\log X_t$ is Gaussian and allows for explicit calculations, see Sections 3.5 and 4.8.

3.2 Snell Envelopes

Instead of directly solving the impulse control problem in (2.4) we shall construct a recursive solution using the concept of Snell envelope for optimal stopping. We

begin with a quick review of the fundamental concepts. Our main references are El Karoui [32] and Karatzas and Shreve [49].

Let Z be an $\mathbb{R}$-valued càdlàg process adapted to the filtration $(\mathcal{F}_t)$ with square-integrable supremum $Z \in \mathscr{S}_T^2$. We emphasize that Z need not be Markov. For a given stopping time ν, define $\mathcal{S}_\nu = \{\tau \leqslant T : \mathbb{F}-\text{stopping time such that}\, \nu \leqslant \tau \text{ a.s. }\}$ to be the set of all stopping times after ν. Thus, $\mathcal{S} \equiv \mathcal{S}_0$ is the set of all $\mathbb{F}$-stopping times bounded by T.

Definition 1. *Let $(X_l, l \in \mathbb{L})$ be an arbitrary family of random variables. The essential supremum of (X_l) is the unique random variable $X = \operatorname{ess\,sup}_l X_l$ such that $X \geqslant X_l$ a.s. for all $l \in \mathbb{L}$, and $X \leqslant Y$ a.s. for all random variables Y such that $Y \geqslant X_l$ a.s. $\forall\, l \in \mathbb{L}$.*

Definition 2. *A family $(X_l, l \in \mathbb{L})$ of random variables is* directed upwards *if for any pair (l', l'') of elements of $\mathbb{L}$, there exists $l \in \mathbb{L}$ such that $X_l \geqslant \sup(X_{l'}, X_{l''})$.*

Armed with the above notions, for each $\nu \in \mathcal{S}$, we now define the nonnegative random variable

$$Y(\nu) \triangleq \operatorname*{ess\,sup}_{\tau \in \mathcal{S}_\nu} \mathbb{E}\Big[Z_\tau \Big| \mathcal{F}_\nu\Big]. \tag{3.5}$$

It is easy to see that for any $\nu \geqslant \tau$

$$\mathbb{E}\big[Y(\nu)|\, \mathcal{F}_\tau\big] = \operatorname*{ess\,sup}_{\mu \in \mathcal{S}_\nu} \mathbb{E}[Z_\mu|\, \mathcal{F}_\tau] \leqslant Y(\tau),$$

so $\{Y(\nu)\}$ forms a supermartingale family. Moreover, this family is directed upwards and is right-continuous in expectation and therefore [32] there exists a càdlàg $\mathbb{F}$-supermartingale Y^r such that $Y(\nu) = Y^r_\nu$, i.e. Y^r stopped at time ν. The process Y^r is called the aggregating supermartingale of the family $\{Y(\nu)\}$. Under the additional assumption that Z is continuous from the left in expectation: $\tau_n \searrow \tau \Longrightarrow \mathbb{E}[Z_{\tau_n}] \to \mathbb{E}[Z_\tau]$, it can now be shown that the minimal optimal stopping time τ^* for $Y(0)$ exists and is explicitly given by

$$\tau^* = \inf\Big\{s \geqslant 0 : Y^r_s \leqslant Z_s\Big\}.$$

That is, $Y_0^r = \mathbb{E}[Z_{\tau^*}] = \sup_\tau \mathbb{E}[Z_\tau]$. Y^r is called the Snell envelope of Z and can be further characterized as the *smallest* càdlàg supermartingale dominating Z. By a basic comparison and Doob's inequality,

$$\mathbb{E}\Big[\sup_{0\leqslant s\leqslant T}|Y_s^r|^2\Big] \leqslant 4\sup_{0\leqslant s\leqslant T}\mathbb{E}\big[(Y_s^r)^2\big] \leqslant 4\,\mathbb{E}[(Y_0^r)^2] \leqslant 4\,\mathbb{E}\Big[\sup_{0\leqslant s\leqslant T}|Z_t|^2\Big] < \infty.$$

We therefore obtain a stability of the set $\mathscr{S}_T^2$ under the operation of Snell envelopes.

To understand the continuity properties of Y^r, apply the Doob–Meyer decomposition, $Y^r = M - A$ for some $\mathbb{F}$-martingale M and $\mathbb{F}$-predictable non-decreasing process A. If Z has continuous paths, then so does A [49, Theorem D.13]. Furthermore, if the filtration $(\mathcal{F}_t)$ is Brownian, then the martingale M and therefore Y^r must also be continuous. More generally,

Proposition 1. *[45, Lemma 1] Suppose Z is upper semicontinuous from the left and of class* [D] *(that is $\{Z_\tau\colon \tau \in \mathcal{S}\}$ is uniformly integrable) and $(\mathcal{F}_t)$ is a completed Brownian filtration. Then Y^r is continuous and of class* [D].

This shows that if we want to preserve the continuity of the Snell envelope, we can only allow positive jumps in the reward process. In any case, on $[0, \tau^*)$ Y^r is a martingale and hence $A_{\tau^*-} = 0$.

3.3 Recasting Optimal Switching as Iterative Optimal Stopping

We return to the optimal switching setting of Section 2.2. To begin solving (2.4) we shall first consider a restricted situation where we put a fixed upper bound on the total number of switches allowed. Define $\mathcal{U}^k(t) \triangleq \{(\xi, \mathcal{T}) \in \mathcal{U}(t)\colon \tau_\ell = T \text{ for } \ell \geqslant k+1\}$ to be the set of all admissible strategies on $[t, T]$ with at most k switches. Denote by $\overline{J^k}$ the value function where we optimize only over $\mathcal{U}^k$,

$$\overline{J^k}(t, X_t, i) = \operatorname*{ess\,sup}_{u\in\mathcal{U}^k(t)} \mathbb{E}\left[\int_t^T \psi_{u_s}(s, X_s)\,ds - \int_t^T C\,|du|_s \,\middle|\, \mathcal{F}_t, u_t = i\right]. \tag{3.6}$$

Since (X_t) is strong Markov the above definition makes sense and can be extended for t being a stopping time. The existence of $\overline{J^k}$ as a measurable left-continuous left-limited process of class [D] is shown by El Karoui [32, Theorem I.21].

It is intuitive that optimization over $\mathcal{U}^{k+1}$ and over $\mathcal{U}^k$ should be related to each other. In fact, by the formal Bellman optimality principle solving the problem with $k+1$ switches is equivalent to finding the optimal first switching time τ which maximizes the initial payoff until τ plus the value function at τ corresponding to optimal switching with k switches.

This train of thought leads us to the method of solving our switching problem through a recursive sequence of simpler optimal stopping problems. More precisely we are going to give an alternative recursive construction for $\overline{J^k}$. Define $J^k(t,x,i)$ where $k = 0, 1, 2, \ldots,\ 0 \leqslant t \leqslant T,\ i \in \mathbb{Z}_M$, via

$$\begin{aligned} J^0(t,x,i) &\triangleq \mathbb{E}\Big[\int_t^T \psi_i(s, X_s)\, ds \Big|\, X_t = x\Big], \\ J^k(t,x,i) &\triangleq \sup_{\tau \in \mathcal{S}_t} \mathbb{E}\Big[\int_t^{T\wedge\tau} \psi_i(s, X_s)\, ds + \mathcal{M}^{k,i}(\tau, X_\tau) \Big|\, X_t = x\Big]. \end{aligned} \tag{3.7}$$

The recursion is hidden inside the operator $\mathcal{M}$ which is known in the literature as the intervention operator. $\mathcal{M}$ specifies the best value that can be achieved by making an immediate switch from current mode i given k switches remaining,

$$\mathcal{M}^{k,i}(t,x) \triangleq \max_{j \neq i} \Big\{ -C_{i,j} + J^{k-1}(t,x,j) \Big\}. \tag{3.8}$$

If there are only two regimes like in Hamadène and Jeanblanc [44], the maximum above is trivial, because there is only one regime to switch into.

We first verify that the definition of J^k is reasonable, i.e. that J^k can be chosen to be $(\Omega \times \mathbb{R}_+ \times E, \mathcal{F} \otimes \mathcal{B}_{\mathbb{R}_+} \otimes \mathcal{B})$-measurable. This is straightforward once we connect to notation of Section 3.2. Pick an initial value $X_0 = x$ and observe that under minimal regularity assumptions (3.7) is equivalent to

$$J^k(t, X_t^x, i) \triangleq \operatorname*{ess\,sup}_{\tau \in \mathcal{S}_t} \mathbb{E}^x \Big[\int_t^\tau \psi_i(s, X_s^x)\, ds + \mathcal{M}^{k,i}(\tau, X_\tau^x) \Big|\, \mathcal{F}_t\Big]. \tag{3.9}$$

Let $\Psi_t^i = \int_0^t \psi_i(s, X_s^x)\, ds$, $Z_t^{k,i} = \int_0^t \psi_i(s, X_s^x)\, ds + \mathcal{M}^{k,i}(t, X_t^x)$. It is easy to see that $Z^{k,i}$ satisfies all the regularity assumptions of Section 3.2. Then $Y^{k,i}$, the Snell envelope of $Z^{k,i}$, solves $Y_t^{k,i} = \operatorname{ess\,sup}_{\tau \in \mathcal{S}_t} \mathbb{E}^x[Z_\tau^{k,i} | \mathcal{F}_t]$ and comparing with (3.9) we see that

$$J^k(t, X_t^x, i) = Y_t^{k,i} - \Psi_t^i.$$

This identification resolves all the measurability issues and moreover shows that the optimal stopping time τ_k^* corresponding to the Snell envelope defined by $J^k(t, x, i)$ is simply $\tau_k^* = \inf\{s \geqslant t \colon J^k(s, X_s, i) = \mathcal{M}^{k,i}(s, X_s)\} \wedge T$. Note that to ease on notation we have labeled by J^k both the raw essential supremum and the aggregating supermartingale of the Snell envelope, cf. (3.5).

Theorem 1. *J^k is equal to the value function for the optimal switching problem with at most k switches allowed, $\overline{J^k}$.*

Proof. The theorem states that a 'global' optimization using a fixed maximum of k switches is equivalent to a successive 'local' optimization using one switch at a time in (3.7).

Our proof is based on direct use of the properties of Snell envelope. We induct on the number of switches left. The case $k = 1$ is immediate. Indeed, $\overline{J^1}$ is the value function of a simple optimal stopping problem (all we can choose is (ξ_1, τ_1)) and similarly J^1 is directly the Snell envelope of

$$\operatorname*{ess\,sup}_{\tau, j} \mathbb{E}\Big[\int_t^\tau \psi_i(s, X_s)\, ds - C|i - j| + \int_\tau^T \psi_j(X_s)\, ds \,\Big|\, \mathcal{F}_t\Big].$$

Next define $\tau_0^* = 0$, and for $\ell = 1, \dots, k$ the stopping times

$$\tau_\ell^* \triangleq \inf\Big\{s \geq \tau_{\ell-1}^* \colon J^\ell(s, X_s, i) = \max_{j \neq i}\big(-C_{i,j} + J^{\ell-1}(s, X_s, j)\big)\Big\} \wedge T, \tag{3.10}$$

and sequence of regimes $\xi_\ell^* \triangleq \arg\max_j \mathcal{M}^{\ell,i}(\tau_\ell^* -, X_{\tau_\ell^* -})$. It is easy to see that by 'unrolling' the value functions J^k using the above (ξ^*, τ^*) we obtain a well-defined strategy $u^* \in \mathcal{U}^k(t)$ for optimal switching with at most k switches. To show that this

strategy dominates any other, let $u = (\xi_1, \xi_2, \dots;\ \tau_1, \tau_2, \dots, \tau_k)$ be a given strategy in $\mathcal{U}^k(t)$. Then by construction

$$\begin{aligned} J^k(t,x,i) &\geqslant \mathbb{E}\Big[\int_t^{\tau_1} \psi_i(s, X_s^{t,x})\, ds + \mathcal{M}^{k,i}(\tau_1, X_{\tau_1}^{t,x})\Big|\, X_t = x\Big] \\ &\geqslant \mathbb{E}\Big[\int_t^{\tau_1} \psi_i(s, X_s^{t,x})\, ds - C|i-\xi_1| + J^{k-1}(\tau_1, X_{\tau_1}^{t,x}, \xi_1)\Big|\, X_t = x\Big], \end{aligned} \tag{3.11}$$

but $\mathbb{E}[J^{k-1}(\tau_1, X_{\tau_1}, \xi_1)|\, \mathcal{F}_t] = \mathbb{E}[\overline{J^{k-1}}(\tau_1, X_{\tau_1}, \xi_1)|\, \mathcal{F}_t]$ a.s. by induction. Comparing with (3.6), $J^k \geqslant \overline{J^k}(\cdot; u)$ and since u was arbitrary, $J^k \geqslant \overline{J^k}$. We also see that the inequalities become equalities for (ξ_1^*, τ_1^*) because τ_1^* is the optimal time from the Snell envelope characterization and $J^k(s, X_s^{t,x}, i)$ is a martingale on $[t, \tau_1^*]$. Therefore $\overline{J^k}(\cdot; u^*)$ achieves the supremum in the impulse control problem (3.6). Observe that as a corollary we obtain the *existence* of an optimal switching policy and a direct proof that this policy is of the 'barrier' type—the optimal switching times are hitting times for (X_t). □

As the next theorem demonstrates, taking the limit $k \to \infty$ in the number of switches we recover the true value function.

Theorem 2. *Define* $J(t,x,i)$ *by* (2.4) *and* $J^k(t,x,i)$ *by* (3.7). *Then* $\lim_{k\to\infty} J^k = J$ *pointwise.*

Proof. Since having more switches is always advantageous, $J^{k+1} \geqslant J^k$. At the same time,

$$J^k(t,x,j) \leqslant \mathbb{E}^x\Big[\int_0^T \max_i |\psi_i(s, X_s)|\, ds\Big] < \infty.$$

Therefore, the bounded monotone sequence J^k converges to some finite J^∞. It remains to show that $J^\infty = J$. Let $u^* = (\xi_1^*, \dots; \tau_1^*, \dots)$ be an optimal policy with unlimited number of switches. Define $u_k^*(t) = u^*(t)$ if $t \leqslant \tau_k^*$, and $u_k^*(t) = u^*(\tau_k^*)$ otherwise. That is, u_k^* approximates u^* up to the latter's k-th switch and remains

constant afterwards. Trivially, $J^\infty(t,x,i) \geqslant J(t,x,i;u_k^*)$. To bound the remainder, observe that since a switch is made only if it increases expected value,

$$0 \leqslant \mathbb{E}\Big[\int_{\tau_k^*}^T \psi_{u_s^*}(s,X_s)\,ds - \int_{\tau_k^*}^T C|du^*|_s\Big] \leqslant \mathbb{E}[T-\tau_k^*]\cdot\mathbb{E}\big[\sup_{0\leqslant t\leqslant T}\max_i|\psi_i(t,X_t)|\,\big].$$

Because by admissibility $\tau_k^* \to T$ a.s., the right hand side converges to zero. Hence, $J^\infty(t,x,i) \geqslant \lim_{k\to\infty} J(t,x,i;u_k^*) = \sup_{u\in\mathcal{U}(t)} J(t,x,i;u)$. This also shows that for any $\epsilon > 0$, there is a k large enough such that the optimal control of J^k generates an ϵ-optimal strategy, a fact which is useful for numerical approximations. □

Thanks to our recursive construction, it becomes clear that because (X_t) is Markov, then the optimal policy can be chosen to be Markovian as well [53]. Indeed, each switch by itself becomes Markovian in (3.10). In turn this implies that the Dynamic Programming principle holds for the value function, i.e. with obvious notation

$$J(t,x,i) = \sup_{u\in\mathcal{U}(t,t')} \mathbb{E}\big[H(x,i,[t,t'];u) + J(t',X_{t'}^{t,x},u_{t'})\big|\,X_t = x, u_t = i\big]. \tag{3.12}$$

The idea of representing impulse control problems as limits of sequential optimal stopping has been well known for a long time, see e.g. [53, 62]. In the setting of optimal switching some of the ideas appear in [44], however we believe our work is the first careful formulation in purely probabilistic terms.

3.4 Exponential Maturity Randomization

In the previous section we have replaced a seemingly harder problem of optimal switching with a sequence of simpler optimal stopping problems. We now take a second step of replacing the harder problem with finite horizon by a sequence of infinite horizon ones. Combined, the two methods allow us to give a unified treatment of the general problem of optimal switching as an iterative system of optimal stopping problems on infinite horizon. This conceptual reduction paves the way for new qualitative results regarding the resulting value function.

We introduce the maturity randomization method pioneered by Carr [18] and recently put on firm ground by Bouchard et al. [10]. Instead of finishing at a fixed time T, the contract is terminated at the N-th *arrival*, the time $\bar{T}_N$. Let σ_n denote the inter-arrival times which are i.i.d. exponential random variables with mean $1/\lambda \triangleq T/N$. Following the notation in [10] we count the σ's backwards. Then $\bar{T}_N = \sum_{n=1}^{N} \sigma_n$, and the new problem consists in optimizing over $[0, \bar{T}_N]$. By the Law of Large Numbers in the limit $N \to \infty$, $\bar{T}_N \to T$ in L^2 converging to the finite horizon case.

The key feature of the construction is that the arrivals are observable but independent of (X_t). To achieve this we redefine the big σ-algebra to be $\mathcal{F} = \sigma(\mathcal{F}_\infty \vee I)$ where I represents additional randomness. The inter-arrival times σ_i will be drawn in an i.i.d. fashion from I. We have $I \perp\!\!\!\perp \mathcal{F}_\infty$ and letting $T_n = \sum_{i=N-n}^{N} \sigma_i$ denote the n-th arrival time, the observable filtration is changed to $\tilde{\mathcal{F}}_t = \sigma\big(\mathcal{F}_t \bigvee_n \mathbb{1}_{\{T_n \leqslant t\}}\big)$. Hence, at time t we know exactly how many arrivals have occurred. However, by the memoryless property of exponentials, on each interval $[T_n, T_{n+1}]$ we have time *stationarity*. Overall, we have replaced the smooth but time-dependent switching boundary from the finite horizon case by N constant boundaries, one for each $[T_n, T_{n+1}]$. This major simplification is the raison d'être of maturity randomization.

To make everything precise we make explicit the (possibly random) horizon of the problem, writing $J(t, x, i; T)$ for our old value function. Following [10] we now define a new set of value functions as follows. Fix total number of arrivals N and let $\mathcal{U}^k(t, v)$ be the set of all policies u that use at most k *total* switches and that agree with policy v up to time t: $u_s = v_s$, for $s \leqslant t$. Note that past history matters now, and if we used up switches before, we have fewer left. For a strategy $u \in \mathcal{U}^k(0)$, let

$$V_0^u(t, x_0, i) \triangleq H(x_0, i; [0, t]; u) = \int_0^t \psi_{u_i}(X_s)\, ds - C \int_0^t |du|_s \ \in \tilde{\mathcal{F}}_t, \quad \text{cf. (2.2)},$$

and define recursively,

$$V_n^v(t,x_0,i) \triangleq \operatorname*{ess\,sup}_{u\in\mathcal{U}^k(t,v)} \mathbb{E}\big[V_{n-1}^u(t+\sigma_n,x_0,i)\big|\tilde{\mathcal{F}}_t\big] \tag{3.13}$$
$$= \operatorname*{ess\,sup}_{u\in\mathcal{U}^k(t,v)} \mathbb{E}\Big[\int_t^\infty \lambda \mathrm{e}^{-\lambda(s-t)} V_{n-1}^u(s,x_0,i)\,ds\Big|\tilde{\mathcal{F}}_t\Big], \quad n=1,2,\ldots,N.$$

When $t=0$, we just write $V_n(0,x,i)$. The control problem for $V_n^v(t,x_0,i)$ optimizes over the random subinterval $[t,t+\sigma_n]$ given past history of (X_t) up to t and given that there are still n intervals to go. Note that the running payoff is encoded into $V_n^v(t,x_0,i)$ which is the sum of realized profit on $[0,t]$ using the strategy v plus best expected future reward until $t+\sum_{i=1}^n \sigma_i$. The initial condition $X_0=x_0$ remains constant and only the horizon is changing.

To state the main convergence result we make the following assumptions:

Assumption 2 (HU)**.** *Stability of $\mathcal{U}^k$ under bifurcation at deterministic times: For any $A\in\mathcal{F}_t$ and $v_1,v_2\in\mathcal{U}^k$, $v_1=v_2$ on $[0,t) \Longrightarrow v_1\mathbb{1}_A+v_2\mathbb{1}_{A^c}\in\mathcal{U}^k$.*

Assumption 3 (HY)**.** *Possible rewards are always dominated: $\forall u\ H(x,i;[0,t];u)\leqslant \bar{M}_t$, where $\bar{M}$ is a uniformly integrable martingale on $[0,\infty]$.*

Assumption 4 (HV)**.** *There exists an aggregating $(\Omega\times\mathbb{R}_+\times E,\mathcal{F}\otimes\mathcal{B}_{\mathbb{R}_+}\otimes\mathcal{B})$-measurable version of V_n^v.*

Then under $(HU),(HY)$, and (HV) [10],

$$J(0,x,i;\ \bar{T}_N)\leqslant V_N(0,x,i)\leqslant \int_{\mathbb{R}_+^N} J(0,x,i;\ \sigma^1+\sigma^2+\cdots+\sigma^N)\,m(d\sigma) \tag{3.14}$$

with equality in the limit $N\to\infty$. This says that the recursively defined V_N, which consists of piecewise optimization on each interval $[T_n,T_{n+1}]$ is better than direct optimization on $[0,\bar{T}_N]$, but worse than pathwise optimization with respect to each separate realization of the arrival process. The nontrivial part is proving that in the limit the upper and lower bounds converge, squeezing V_N. Note that the theorem does not guarantee any monotonicity assumptions on the sequence V_N, even though

we intuitively expect that V_N is eventually increasing in N, because having more intervals is akin to 'more information' which should help the optimization.

Verifying the necessary hypotheses on (X_t) and admissible policies $\mathcal{U}$ is straightforward in our case of recursive optimal stopping. Indeed, (HU) is immediate for optimal switching and the measurability hypothesis (HV) follows directly from the Snell envelope construction. The uniform integrability (HY) follows if we assume that $\mathbb{E}[\sup_t \max_m |\psi_m(t, X_t)|] < \infty$ on $[0, \infty]$.

In analogy with Section 3.3 we will now provide a second doubly recursive construction in the number of switches k and the number of random time intervals remaining n. To avoid confusion we relabel our value functions as V instead of J. Suppose we are on the subinterval between T_{N-n-1} and T_{N-n}, $0 < n \leqslant N$ so that n arrivals remain before termination. Let ζ be the time of next arrival. Then ζ is an exponential random variable with mean $1/\lambda$, so that λ is interpreted as the rate of 'dropping' to the level below. Define

$$
\begin{aligned}
V^{n,k}(t,x,i) &= \sup_{\tau \in \mathcal{S}_t,\, j} \mathbb{E}\Big[\int_t^{\zeta \wedge \tau} \psi_i(s, X_s)\, ds + V^{n-1,k}(\zeta, X_\zeta, i) \mathbb{1}_{\zeta < \tau} \\
&\qquad + (V^{n,k-1}(\tau, X_\tau, j) - C_{i,j}) \mathbb{1}_{\zeta > \tau} \Big| \, \tilde{\mathcal{F}}_t, X_t = x, u_t = i \Big] \\
&= \sup_{\tau \in \mathcal{S}_t,\, j} \mathbb{E}\Big[\int_t^{\tau} \mathrm{e}^{-\lambda(s-t)} (\psi_i(s, X_s^{t,x}) + \lambda V^{n-1,k}(s, X_s^{t,x}, i))\, ds \\
&\qquad + \mathrm{e}^{-\lambda(\tau - t)} (V^{n,k-1}(\tau, X_\tau^{t,x}, j) - C_{i,j}) \Big| \, X_t = x \Big], \qquad (3.15) \\
V^{n,0}(t,x,i) &= \mathbb{E}\Big[\int_t^{\infty} \mathrm{e}^{-\lambda(s-t)} (\psi_i(s, X_s^{t,x}) + \lambda V^{n-1,0}(s, X_s^{t,x}, i))\, ds \Big| \, X_t = x \Big],
\end{aligned}
$$

and $V^{0,k} \equiv 0$. The equality between two definitions in (3.15) follows by the memoryless property of the exponentials. By analogy with before, we expect that $V^{N,k}$ actually corresponds to the value function for optimal switching with at most k switches allowed and 'end-of-the-world' at $\bar{T}_N$. Due to memorilessness, if we decide to switch at time τ, the clock restarts and we are once again facing the problem of optimal switching with n periods left and $k-1$ switches.

Theorem 3. *The two methods* (3.13) *and* (3.15) *of optimizing over* $\mathcal{U}^k$*, i.e. with at most* k *switches allowed, are equal:* $V_n(0,x,i) = V^{n,k}(0,x,i)$.

Proof. Conceptually this is a straightforward adaptation of the results in [10] for the case of optimal switching, which is only slightly more general than the optimal stopping considered there. The main notational difficulty is that the definition of V_n^v aggregates previous payoffs while $V^{n,k}$ only looks to the future. For $u = (\xi, \mathcal{T}) \in \mathcal{U}^k$ their general relationship is

$$V_n^u(t,x,i) = V^{n,k-k_t}(t, X_t^x, u_t) + H(x,i,[0,t];u),$$

with $k_t = \sum_n \mathbb{1}_{\tau_n < t}$ the (random) number of switches made by u up to t.

The proof is again by induction on k, n. The conditions are trivially satisfied when $n = 0$ or $k = 0$. Suppose the result is true in $\mathcal{U}^{k-1}$ with n periods and in $\mathcal{U}^k$ with $n-1$ periods. We will show it is also true in $\mathcal{U}^k$ with n periods. For any $u \in \mathcal{U}^k(t,v)$ isolate out the first switching decision as

$$u = \Big\{(u_1, u') : u' \in \mathcal{U}^k(\tau_1, v \cup u_1), \quad u_1(s) = u_1(t) = v(t) \text{ for } t \leqslant s \leqslant \tau_1\Big\}.$$

This just means that until τ_1 we do not switch and then use u'. Now conditioning on the order of ζ and τ_1 we have

$$\begin{aligned} V_n^v(t,x,i) &= \operatorname*{ess\,sup}_{u \in \mathcal{U}^k(t,v)} \mathbb{E}\big[V_{n-1}^u(t+\zeta, x, i)\big|\, \tilde{\mathcal{F}}_t\big] \\ &= \operatorname*{ess\,sup}_{u \in \mathcal{U}^k(t,v)} \mathbb{E}\big[V_{n-1}^u(t+\zeta, x, i)\mathbb{1}_{t+\zeta \leqslant \tau_1} + V_{n-1}^u(t+\zeta, x, i)\mathbb{1}_{t+\zeta > \tau_1}\big|\, \tilde{\mathcal{F}}_t\big]. \end{aligned} \tag{3.16}$$

For the second term we use the fact that conditional on $\zeta > \tau_1 - t$, the distribution of ζ is still exponential, due to the independence between I and $\mathcal{F}_\infty$. Directly,

$$\begin{aligned} \mathbb{E}\big[V_{n-1}^u(t+\zeta,x,i)\mathbb{1}_{t+\zeta>\tau_1}\big|\, \tilde{\mathcal{F}}_t\big] &= \mathbb{E}\big[\int_{\tau_1}^\infty \lambda \mathrm{e}^{-\lambda(s-t)} V_{n-1}^{u'}(s,x,i)\, ds\big|\, \tilde{\mathcal{F}}_t\big] \\ &= \mathbb{E}\Big[\mathbb{E}[\int_{\tau_1}^\infty \lambda \mathrm{e}^{-\lambda(s-\tau_1)} V_{n-1}^{u'}(s,x,i)\, ds|\, \tilde{\mathcal{F}}_{\tau_1}] \cdot \mathrm{e}^{-\lambda(\tau_1 - t)}\Big|\tilde{\mathcal{F}}_t\Big] \\ &= \mathbb{E}\big[V_n^{u'}(\tau_1, x, i) \cdot \mathrm{e}^{-\lambda(\tau_1 - t)}\,\big|\,\tilde{\mathcal{F}}_t\big]. \end{aligned} \tag{3.17}$$

The last step is justified by the continuity properties of the control problem for V_n^v which allows (3.13) to be extended for any $\mathbb{F}$-stopping time τ using the standard càdlàg modification arguments [32, p.99]. Since u' uses one more switch than u we can invoke the inductive assumption,

$$\begin{aligned} V_n^{u'}(\tau_1, x, i) &= V^{n,k-k_t-1}(\tau_1, X_{\tau_1}^x, u'_\tau) + H(x, i, [0, \tau_1]; u') \\ &= \big\{V^{n,k-k_t-1}(\tau_1, X_{\tau_1}^x, u'_\tau) - C_{u_t,u'_\tau} + \int_t^{\tau_1} \psi_{u_t}(X_s^x)\, ds\big\} + H(x, i, [0, t]; u). \end{aligned}$$

As for the first term in (3.16), by invoking the inductive assumption with $n' = n-1, k' = k$ we have

$$\begin{aligned} \mathbb{E}\big[V_{n-1}^u(t+\zeta, x, i)\mathbb{1}_{t+\zeta\leqslant\tau_1}\big|\,\tilde{\mathcal{F}}_t\big] &= \mathbb{E}\big[\int_t^{\tau_1} \lambda \mathrm{e}^{-\lambda(s-t)} V_{n-1}^u(s, x, i)\, ds\big|\,\tilde{\mathcal{F}}_t\big] \\ &= \mathbb{E}\Big[\int_t^{\tau_1} \lambda \mathrm{e}^{-\lambda(s-t)} V^{n-1,k-k_t}(s, X_s, u_s)\, ds \\ &\qquad + \int_t^{\tau_1} \psi_{u_t}(X_s)\, \mathrm{e}^{-\lambda(s-t)} ds + H(x, i, [0, t]; u)\big|\,\tilde{\mathcal{F}}_t\Big], \end{aligned} \tag{3.18}$$

where the last term is coming from $\int_t^{\tau_1} \lambda \mathrm{e}^{-\lambda(s-t)}\, H(x, i, [0, s]; u)\, ds$. Roughly speaking, when $t+\zeta < \tau_1$, $V_n^u(t+\zeta)$ is just optimizing in $\mathcal{U}^k(t+\zeta, u)$ with one less arrival left. However, because ζ is independent from (X_t) we cannot apply the dynamic programming principle at instant ζ and must resort to the explicit calculation above. Comparing (3.16),(3.17) and (3.18) with (3.15) the induction is complete. □

By an analogue of Theorem 2, the original value function $J(t, x, i)$ equals to $\lim\limits_{N\to\infty}\lim\limits_{k\to\infty} V^{N,k}(t, x, i)$. At the same time, if we set $R_\lambda f(x) \triangleq \mathbb{E}^x\big[\int_0^\infty \mathrm{e}^{-\lambda t} f(X_t)\, dt\big]$, then (3.15) can be rewritten as

$$\begin{aligned} V^{n,k}(t, x, i) = \sup_{\tau\in\mathcal{S}_t,\, j} \mathbb{E}\Big[\mathrm{e}^{-\lambda(\tau-t)}\big\{V^{n,k-1}(\tau, X_\tau^{t,x}, j) - C_{i,j} \\ - R_\lambda\Big(\psi_i(s, X_\tau^{t,x}) + \lambda V^{n-1,k}(\tau, X_\tau^{t,x}, i)\Big)\big\}\big|\, X_t = x\Big] + R_\lambda\big(\psi_i + \lambda V^{n-1,k}(0, \cdot, i)\big)(x) \end{aligned}$$

making it a standard optimal stopping problem for a stationary Markov process. Hence, the entire well-developed theory for this field can be brought to bear. Es-

pecially interesting is the connection between Snell envelopes and minimal excessive majorants that we explore in in Section 3.6.1.

3.5 Regularity of the Value Functions

In this section we state several results on the regularity of $J(t,x,i)$. Even though the required conditions are increasingly restrictive, they still cover the situations of practical interest. For a typical (X_t) think of an exponential of a linear diffusion, either mean-reverting OU or Brownian motion with drift. A typical payoff rate would be linear and increasing in the components of X_t.

Theorem 4. *Suppose that (X_t) is an d-dimensional OU process of (3.4) and the payoff rates $\psi_i(t,x)$ are convex. Then the value functions $J(t,x,i)$ are convex in x.*

Proof. The key property we need is the linearity of the OU process,

$$X_s^{x+\epsilon e_i} = X_s^x + \mathrm{e}^{-\kappa^i s}\epsilon e_i, \quad \text{where } e_i = (0,\dots,0,1,0,\dots), \tag{3.19}$$

with a one in the i-th position. We first prove the result assuming the payoff rates are linear, whence we write them as $\psi_i(s,X_s) = A_i \cdot X_s - \psi(s,i)$ for some set of (time-dependent) vectors (A_i). Fix initial condition (x,i) and let $\hat{u}^x$ be the respective optimal strategy in $\mathcal{U}(t)$,

$$J(t,x,i) = \mathbb{E}\Big[\int_t^T [A_{\hat{u}_s^x}\cdot X_s^x - \psi(s,\hat{u}_s^x)]\,ds - \int_t^T C|d\hat{u}^x|_s\Big|\, X_t = x\Big]. \tag{3.20}$$

Without loss of generality we concentrate on the first component $X^{(1)}$ of (X_t). If we perturb the initial condition in the $X^{(1)}$ direction via $X_0 = x+\epsilon e_1$, $\hat{u}^x$ is still an acceptable strategy to use, so that

$$\begin{aligned} J(t,x+\epsilon e_1,i) &\geqslant \mathbb{E}\Big[\int_t^T [A_{\hat{u}_s^x}\cdot X_s^{x+\epsilon e_1} - \psi(s,\hat{u}_s^x)]\,ds - \int_t^T C|d\hat{u}^x|_s\Big|\, X_t = x+\epsilon e_1\Big] \\ &= J(t,x,i) + \epsilon\,\partial_1(t,x_1,i) \end{aligned}$$

for some function ∂_1 due to (3.19) and (3.20). Now in reverse, starting with $x + \epsilon e_1$ and looking at $\hat{u}^{x+\epsilon e_1}$ we obtain

$$J(t, x, i) \geqslant J(t, x + \epsilon e_1, i) - \epsilon\, \partial_1(t, x_1 + \epsilon, i),$$

which must imply that fixing the other parameters $\partial_1(t, x, i)$ is increasing in x. Now letting $\epsilon \to 0$, we see that $\partial_1(t, x, i)$ is nothing but the right derivative in the first coordinate direction $\partial J(t, x, i)/\partial x_1^+$. So we showed that the derivative of the value function is increasing in x, i.e. $J(t, x, i)$ is convex in x.

For a general convex ψ_i, write it as a supremum of piecewise linear functions. For any truncated piecewise linear approximation $\psi_i(x) \simeq \max_{1 \leqslant j \leqslant M} \psi_{ij}(x)$, enlarge the policy space by replacing regime i with a collection $\{i_1, i_2, \ldots i_M\}$, such that the payoff rate at regime i_j is ψ_{ij} and the switching cost between two regimes in the collection is zero. Once we allow as many of those zero-cost switches as desired, the acceptable strategies reduce to the set $\mathcal{U}$ and we have an obvious correspondence to the original setting. The first part of the proof applies now to the extended policies and we are done. □

A related result in the special case of American options and one-dimensional (X_t) appeared in El Karoui et al. [34]. There it is also shown that if one does not immediately switch at (t, x) then the left-derivative of $J(t, x, i)$ is bounded by the left-derivative of $\psi_i(x)$.

We now suppose that the flow of (X_t) is Lipschitz, namely that μ and σ in (3.1) are continuously differentiable with bounded derivatives. By a standard result this implies $\mathbb{E}^x|X_t^x|^2 \leqslant C(1 + \|x\|^2)$, or more generally, $\mathbb{E}^x[\sup_t \|X_t\|^p] \leqslant C(1 + \|x\|^p)$ [65, p. 342]. As a corollary we obtain the following two lemmas.

Lemma 1. *Suppose all the payoff rates $\psi_i(t, x)$ are Lipschitz in x. Then the value functions $J(t, x, i)$ are Lipschitz in x.*

Proof. This is a simple estimate :

$$\begin{aligned}|J(t,x,i) - J(t,y,i)|^2 &\leqslant \sup_{u\in\mathcal{U}(t)} \mathbb{E}\big[\int_t^T |\psi_{u_s}(s, X_s^{t,x}) - \psi_{u_s}(s, X_s^{t,y})|^2\, ds\big] \\ &\leqslant C\,\mathbb{E}\Big[\int_t^T |X_s^{t,x} - X_s^{t,y}|^2\, ds\Big] \\ &\leqslant C\,\mathbb{E}\big[\sup_{t\leqslant s\leqslant T} |X_s^{t,x} - X_s^{t,y}|^2\big] \leqslant C\|x-y\|^2,\end{aligned}$$

where C denotes a generic constant that changes from expression to expression. □

Lemma 2. *Suppose the payoff rates ψ_i and the process (X_t) are time-homogeneous and $\mathbb{E}[\sup_s |\psi_i(X_s)|] < \infty$ for all i. Then the value function $J(t,x,i)$ is Lipschitz continuous in t.*

Proof. Because of the time-homogeneity, $J(t,x,i;T) = J(0,x,i;T-t)$, i.e. we can shift the problem to time zero and maturity $(T-t)$. Therefore for $t < t'$,

$$\begin{aligned}|J(t,x,i) - J(t',x,i)| &= |J(0,x,i;T-t) - J(0,x,i;T-t')| \\ &\leqslant \sup_{u\in\mathcal{U}(T-t')} \mathbb{E}^x\Big[\int_{T-t'}^{T-t} |\psi_{u_s}(X_s^x)|\, ds\Big] \\ &\leqslant (t'-t)\,\mathbb{E}\big[\sup_s \max_i |\psi_i(X_s^x)|\big],\end{aligned}$$

and the last expression is uniform in $(t'-t)$ by Assumption 1. □

Theorem 5. *Suppose (X_t) is time-homogeneous, the costs are additive: $C_{i,j} = C_{i,\ell} + C_{\ell,j}$ whenever $i < \ell < j$, and coordinate-wise monotone in payoff $\partial_{x_n}(\psi_i - \psi_j) \geqslant 0$ (alternatively $\leqslant 0$) for all $i > j$. Then the optimal policy for $J(t,x,i)$ can be characterized by exercise boundaries that divide the base space E into connected regions.*

Proof. Additive costs imply that $C_{i,j} = C|i-j|$ which is the case we originally considered in Section 2.2. Consider now two optimal policies $\hat{u}$ and $\hat{v}$ corresponding to starting with same initial condition $X_0 = x$ but different initial regimes. We claim that if $\hat{u}_0 > \hat{v}_0$ then $\hat{u}_t \geq \hat{v}_t\ \forall t$. Clearly, along any given path of (X_t), whenever

$\hat{u}_s = \hat{v}_s$ then the optimal policies coalesce after s by the Markov property. For the order to be reversed, there must therefore be some time τ such that $\hat{u}_{\tau-} > \hat{v}_{\tau-}$ but $\hat{u}_\tau < \hat{v}_\tau$. Without loss of generality assume that τ is a switching time of $\hat{v}$ with $\hat{v}_{\tau-} = 0, \hat{u}_{\tau-} = \hat{u}_\tau = 1, \hat{v}_\tau = 2$. Then being in regime 0 at τ we must prefer regime 2 to regime 1, i.e. $J(\tau, X_\tau, 2) - 2\,C > J(\tau, X_\tau, 1) - C$. But then $J(\tau, X_\tau, 1) < J(\tau, X_\tau, 2) - C$ and so at τ one should also switch from regime 1 to regime 2, contradicting $\hat{u}_\tau = 1$.

The claim is proven and by assumption $\partial_x(\psi_{\hat{u}_t} - \psi_{\hat{v}_t}) \geqslant 0$ ($\leqslant 0$). Therefore the same is true of the respective value functions. It follows that the difference between any two $J(t, x, i)$ and $J(t, x, j)$ is always strictly increasing (strictly decreasing) and the set

$$Switch(i, j; t) \triangleq \Big\{ x \in \mathbb{R}^d : -C_{i,j} + J(t, x, j) > \max_{j' \neq j} \big\{ -C_{i,j'} + J(t, x, j') \big\} \Big\}$$

of all x's where it is optimal to immediately switch from regime i to j is connected. The switching boundaries are the boundaries of $Switch(i, j; t)$ as a function of t for different pairs of (i, j). By Lemma 2 the value function J is continuous in t, and consequently so is the switching set and its boundary. □

3.6 Review of the Variational Formulation

We recall that optimal switching is a special case of impulse control. Redefine the controlled (Markov) state process to be $\tilde{X}_t = (X_t, u_t)$ with values in $\mathbb{R}^d \times \mathbb{Z}_M$. Applying the impulse (ξ, τ) costs $C(\tilde{X}_{\tau-}, \xi)$ and changes the regime only, $\tilde{X}_\tau = (X_{\tau-}, \xi)$. In general, $\tilde{X}_\tau = \tilde{X}_{\tau-} + \xi$ and the aim of the controller is to maximize the discounted future reward up to final time T_0, which is either a killing time of $\tilde{X}$ or an exogeneous problem horizon,

$$J(t, x) = \sup_{(\xi, \mathcal{T}) \in \mathcal{U}(t)} \mathbb{E}\Big[\int_t^{T_0} \mathrm{e}^{-rs} \Psi(\tilde{X}_s)\, ds - \sum_{\tau_k < T_0} \mathrm{e}^{-r\tau_k} C(\tilde{X}_{\tau_k -}, \xi_k) \Big| \, \tilde{X}_t = x \Big].$$

Such problems have been looked at extensively in finance, e.g. for optimal trading with fixed transaction costs [51] or foreign exchange intervention [47].

The classical method of solving impulse control problems driven by Markov processes goes back to the fundamental work of Bensoussan and Lions [8]. The idea is to study the parabolic partial differential equation resulting from applying Bellman's principle to the process conditional on no control on $[t, t+dt)$, together with the equation corresponding to applying optimal control at t. At least one of these two equations must hold at any given instant. Combined we obtain a (quasi-) variational formulation with a free boundary.

The key connection to analysis is furnished by Dynkin's formula:

Proposition 2. *[61, Theorem 7.4.1] Let (X_t) be an Itô diffusion with space-time generator*

$$\mathcal{L}_X \triangleq \partial_t + \sum_i \mu_i(x)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j}\big(\sigma(x)\sigma(x)^T\big)_{ij}\,\frac{\partial^2}{\partial x_i\,\partial x_j}$$

and $f \in \mathcal{C}^{1,2}([0,T]\times E)$. Suppose τ is a stopping time, $\mathbb{E}^x[\tau] < \infty$ and $f(t, X_t)$ is bounded on $[0,\tau)$. Then

$$\mathbb{E}^x[f(\tau, X_\tau)] = f(0,x) + \mathbb{E}^x\Big[\int_0^\tau \mathcal{L}_X f(s, X_s)\,ds\Big]. \tag{3.21}$$

The main tool now is the verification theorem. It states that a smooth solution of the quasi-variational inequality (QVI) constructed below is in fact the value function of the impulse control problem. We adapt it to optimal switching where the impulses affect only the operating regime.

Proposition 3 ([12, 62])**.** *Let X be a Markov càdlàg process on $E \subseteq \mathbb{R}^d$ with space-time generator $\mathcal{L}_X$. Denote by $\mathcal{M}\phi(t,x,i) = \max_{j\neq i}\{-C_{i,j} + \phi(t,x,j)\}$ the intervention operator. We again think of $\mathcal{M}$ as the barrier above which our solution must remain. Let*

$$D = \cup_i\big\{(t,x)\colon \phi(t,x,i) = \mathcal{M}\phi(t,x,i)\big\}$$

be the free boundary. Suppose there exists $\phi(t,x,i)$ *belonging to* $\mathcal{C}^{1,2}\big(([0,T_0]\times E)\setminus D\big)\cap \mathcal{C}^{1,1}(D)$ *and satisfying the following quasi-variational Hamilton-Jacobi-Bellman (HJB) inequality for all* $i\in\mathbb{Z}_M$*:*

i. $\phi \geqslant \mathcal{M}\phi$,

ii. $\mathbb{E}^x\big[\int_0^T \mathbb{1}_{\phi\leqslant\mathcal{M}\phi}\,dt\big]=0$,

iii. $\mathcal{L}_X\phi(t,x,i)+\psi_i(t,x)\leqslant 0, \qquad \phi(T,x,i)=0$,

iv. $\Big(\mathcal{L}_X\phi(t,x,i)+\psi_i(t,x)\Big)\Big(\phi(t,x,i)-\mathcal{M}\phi(t,x,i)\Big)=0.$

Then ϕ *is the optimal value function for the switching problem* (2.4).

In words, the conditions on ϕ require it to dominate the barrier and satisfy the fundamental PDE inside the continuation region. In addition, the free boundary must be such that the process spends Lebesgue-measure zero time on it on the entire interval $[0,T_0]$[2] .

The proof relies on application of (3.21). Since ϕ is potentially unbounded, let τ_R be the minimum of T_0 and the first exit time for (X_t) from a ball of radius R. Then for any stopping time τ,

$$\begin{aligned}\phi(t,x,i) &= \mathbb{E}\Big[-\int_t^{\tau\wedge\tau_R}\mathcal{L}_X\phi(s,X_s^{t,x},i)\,ds+\phi(\tau\wedge\tau_R,X_{\tau\wedge\tau_R}^{t,x},i)\Big|\,X_t=x\Big]\\ &\geqslant \mathbb{E}\Big[\int_t^{\tau\wedge\tau_R}\psi_i(s,X_s^{t,x})\,ds+\phi(\tau\wedge\tau_R,X_{\tau\wedge\tau_R}^{t,x},i)\Big|\,X_t=x\Big], \qquad (3.22)\end{aligned}$$

where we have plugged in condition (iii). Next we let the intervention times be the hitting times of the barrier $\tau_k \triangleq \inf\{s>\tau_{k-1}\colon \phi(s,X_s,u_{\tau_{k-1}})\leqslant\mathcal{M}\phi(s,X_s,u_{\tau_{k-1}})\}\wedge T$, and ξ_k be such that $\phi(\tau_k-,X_{\tau_k-},\xi_k)=\mathcal{M}\phi(\tau_k-,X_{\tau_k-},\xi_{k-1})$. Equality now holds in (iii) and therefore in (3.22) on each subinterval $[\tau_k,\tau_{k+1})$. By induction it can then

[2]In the literature an alternative notation is to say that ϕ is stochastically-$\mathcal{C}^2$ on E.

be shown [8, Lemma 6.3.8] that assuming $\tau_k < T \wedge \tau_R$, and calling u^* the strategy of using (τ_k, ξ_k) above,

$$\phi(0,x,i) = \mathbb{E}\Big[\int_0^{\tau_k} \psi_{u_t^*}(t, X_t)\, dt + \phi(\tau_k, X_{\tau_k}, u^*_{\tau_k}) \\ - C\sum_{i=1}^{k} |u^*_{\tau_i} - u^*_{\tau_i -}| \;\Big|\; X_0 = x, u_0 = i\Big], \quad (3.23)$$

with inequality for any other strategy u. The rest of the proof is like the proof of Theorem 2 in Section 3.3.

While providing a very general framework, the variational approach has a large overhead. Analysis of the QVI without a priori assumption of a smooth solution is quite involved. Besides classical strong solutions, there are all sorts of weak solutions. For example, it can be shown [8, Chapter 6] that the increasing sequence of solutions ϕ^k which consists in applying the optimal control up to τ_k, cf. (3.23) converges upwards to the *minimum* solution of the QVI. At the same time, the sequence of approximate solutions resulting from solving the problem using at most k controls, cf. (3.6) converges to the *maximum* solution of the QVI. As the names suggest, the two are not necessarily equal. To achieve uniqueness, one must pass to the notion of viscosity solutions. Fortunately, viscosity solutions are naturally compatible with stochastic control. Thus, the value function J is always the (unique) viscosity solution of the system in Proposition 3, furnishing the converse to the verification theorem above. Viscosity solutions are beyond the scope of this thesis and we refer to the volume of Fleming and Soner [40] for extensive discussion in the context of general stochastic control.

For implementation, the 'smooth pasting' condition of being $\mathcal{C}^1$ across D is crucial for determining the free boundary. However, it is often assumed without justification, by invoking heuristic arguments. Even then it only leads to a system of implicit equations with the existence of a solution again unclear. See for example the paper of Brekke and Øksendal [12, p. 1031], where the free boundary is parametrized by four

quantities $a_1, \ldots, a_4$ and a complicated coupled algebraic system of equations in a_j is obtained. The authors then state "we have not been able to derive tractable general conditions for existence of a solution" to this system. To sum up, the variational method poses a multitude of technical challenges that prevent rigorous solutions in all but the simplest settings. It is a theoretical overkill that underperforms in practice.

3.6.1 Smallest Excessive Majorants

The original probabilistic treatment of optimal stopping is due to Dynkin [30]. Recall,

Definition 3. *A nonnegative function $f \in \mathcal{C}^1(E)$ is called* β-excessive *if for all $x \in E$,*

$$f(x) \geqslant \mathbb{E}^x[\mathrm{e}^{-\beta t} f(X_t)] \quad \textit{for all } \ t > 0, \ \textit{ and } \quad \lim_{t \searrow 0} \mathbb{E}^x[\mathrm{e}^{-\beta t} f(X_t)] = f(x).$$

If $\beta = 0$ we just call f excessive.

Dynkin's approach is based on the fact that the Snell envelope is the smallest excessive majorant (s.e.m.) of the payoff function. S.e.m.'s may be sometimes found directly by inspection. For example, in the special case when (X_t) is a standard one-dimensional Brownian motion, excessive functions are characterized as concave functions. Recently this result has been extended by Dayanik [24] to all regular one-dimensional diffusions. In general, to find the s.e.m. Q of f, let $(Q_n f)(x) = \max_k \big(f(x), \mathbb{E}^x[f(X_{2^{-n}})], \ldots, \mathbb{E}^x[f(X_{k \cdot 2^{-n}})], \ldots\big)$. Then $Q(x) = \lim_{n \to \infty} Q_n f(x)$. [66, Lemma III.1] Observe how Q_n essentially discretizes the set of stopping times to be multiples of 2^{-n}.

Excessive functions are fundamentally smooth. If (X_t) is any càdlàg strong Markov process and f is excessive for (X_t), then f is continuous [31, Theorem 12.4]. Moreover, if (X_t) is continuous then excessive functions are superharmonic for the

generator $\mathcal{L}_X$, i.e. $\mathcal{L}_X f \leqslant 0$ in the variational sense. This is an extension of the smoothness properties of convex functions which satisfy "$\Delta f \leqslant 0$".

Corollary 1. *Suppose (X_t) is a càdlàg strong Markov process on $E \subseteq \mathbb{R}^d$ and all the payoff rates are time-homogeneous. Then the value function $V^{n,k;N}$ coming from the exponential maturity randomization* (3.15) *is N/T-excessive.*

The above proposition follows more or less directly from [66, Theorem III.3.1] since as already stated the problem for $V^{n,k}$ is a standard optimal stopping problem for a stationary process discounted at rate $\lambda = N/T$. Now the original value function J is the limit as $N \to \infty$, $\lambda \to \infty, k \to \infty$ of $V^{N,k}$. This provides an interesting angle of attack for studying the smoothness of J. However, one must be careful since the excessivity property is non-uniform and depends on total number of exponential periods N.

In the one-dimensional case we have a more precise condition recently found by Dayanik and Egami [23].

Definition 4. *Set $F^\lambda = \psi^\lambda/\phi^\lambda$ where ψ^λ (resp. ϕ^λ) is the increasing (decreasing) fundamental solution of $\mathcal{L}_X u - \lambda u = 0$. A function $u : E \to \mathbb{R}$ is called F^λ-concave if for every $l \leqslant x \leqslant r$*

$$u(x) \geqslant u(l)\frac{F(r) - F(x)}{F(r) - F(l)} + u(r)\frac{F(x) - F(l)}{F(r) - F(l)}.$$

Proposition 4 ([23], Theorem 5.1)**.** *Suppose (X_t) is a regular one-dimensional diffusion with domain E an open subset of $\mathbb{R}$. Then the value function of the impulse control problem*

$$V_1(x,i) \triangleq \sup_{u \in \mathcal{U}} \mathbb{E}\big[H(x,i,[0,\bar{T}_1];u)\big], \qquad \bar{T}_1 \sim exp(\lambda)$$

is F^λ-concave on $int(E)$.

The significance of this proposition is that there are simple geometric methods for finding F^λ-concave majorants. Consequently, the above result shows that there

are potentially even closed form solutions to the optimal switching problem with an exponential horizon. It would be interesting to further explore this aspect using the Dynkin-Dayanik techniques.

3.7 Reflected Backward SDEs

The reflected backward stochastic differential equations (BSDE) are the probabilistic counterpart of the variational approach when (X_t) is an Itô diffusion. For a given initial condition $X_0 = x$, suppose $\exists (Y^x, Z^x, A)$[3] adapted to $(\mathcal{F}_t^X)$ such that $\mathbb{E}\big[\sup_{0\leqslant t\leqslant T} |Y_t^x|^2 + \int_0^T \|Z_t^x\|^2\, dt + |A_T|^2\big] < \infty$, A is continuous and increasing and

$$
\begin{aligned}
Y_t^x &= \int_t^T \psi_i(s, X_s^x)\, ds + A_T - A_t - \int_t^T Z_s \cdot\, dW_s, \qquad (3.24)\\
Y_t^x &\geqslant \mathcal{M}^{k,i}(t, X_t^x), \qquad \text{as defined in (3.8)}\\
\int_0^T &(Y_t^x - \mathcal{M}^{k,i}(t, X_t^x))\, dA_t = 0, \qquad A_0 = 0.
\end{aligned}
$$

The interpretation is that Z is a conditional expectation process that helps Y_t to be $\mathcal{F}_t$-measurable, while A is a compensator that increases only when Y hits the barrier $\mathcal{M}^{k,i}$. Let us also mention that the integrability assumption $\mathcal{M}^{k,i} \in \mathscr{S}_T^2$ that we checked in Theorem 1 is necessary for (3.24) to make sense. The first result in [35, Prop. 2.3] shows that $Y_0^x = J^k(0, x, i)$. More generally, $Y_t^x = J^k(t, X_t^x, i)$. For the intuition of why this might be true, it suffices to observe that by construction of (3.24) for any stopping time τ,

$$
\begin{aligned}
Y_t^x &= \mathbb{E}\Big[\int_t^\tau \psi_i(s, X_s)\, ds + Y_\tau^x + A_\tau - A_t \Big|\, \mathcal{F}_t\Big]\\
&\geqslant \mathbb{E}\Big[\int_t^\tau \psi_i(s, X_s)\, ds + \mathcal{M}^{k,i}(\tau, X_\tau)\mathbb{1}_{\tau<T}\Big]
\end{aligned}
$$

and $A_t = 0$ until Y hits the barrier. It follows that equality above holds for $\tau = \tau^* = \inf\{s > 0 \colon Y_s^x = \mathcal{M}^{k,i}(s, X_s)\}$ which must be optimal.

[3] Of course, all the parameters are functions of initial mode i as well

Let us next state some estimates on Y which will be especially important for our convergence proofs in Section 4.3. We work with a slightly modified version where the barrier is an explicit function of X_t:

$$\begin{cases} Y_t = \int_t^T \Psi(s, X_s)\, ds + A_T - A_t - \int_t^T Z_s \cdot dW_s, \\ Y_t \geqslant g(t, X_t), \qquad \text{and} \qquad \int_0^T (Y_t - g(t, X_t))\, dA_t = 0. \end{cases} \tag{3.25}$$

Assuming Ψ and g are uniformly Lipschitz in x and t,

$$|\Psi(t,x) - \Psi(t',x')| + |g(t,x) - g(t',x')| \leqslant C(1 + |t - t'| + \|x - x'\|_\infty)$$

we have global estimates [35]

$$\mathbb{E}[\sup_{0 \leqslant t \leqslant T} |Y_t|^2 + A_T^2] \leqslant C\, \mathbb{E}\Big[\int_0^T |\Psi(t, X_t)|^2 dt + \sup_{0 \leqslant t \leqslant T} |g(t, X_t)|^2\Big] \tag{3.26}$$

$$\text{or} \qquad \big\| \sup_{0 \leqslant t \leqslant T} |Y_t| \big\|_p \leqslant C\big(1 + \| \sup_{0 \leqslant t \leqslant T} |X_t| \|_p\big).$$

Secondly, (3.25) satisfies a stability property: if Y' solves the same system but with modified barrier g' and generator Ψ' then

$$|Y_t - Y_t'| \leqslant \mathrm{e}^{C(T-t)}(\|g - g'\|_\infty + (T-t)\|\Psi - \Psi'\|_\infty).$$

Note that the last equation implies uniqueness of solutions to (3.25).

The BSDE formulation clarifies the continuity properties of the Snell envelope. Indeed, if the barrier $\mathcal{M}^{k,i}$ is continuous in t, it is clear that the solution Y of (3.24) is also continuous. Furthermore, analogues of stochastic flow theory imply that under regularity conditions on the parameters of the stochastic differential equation for (X_t), Y^x will be continuously differentiable in the initial condition x and other parameters of the (X_t)-SDE [36]. By induction the same is true of the recursive value functions $J^k(t, x, i)$ of our problem (but not necessarily of $J(t, x, i)$ itself).

The link between BSDEs and parabolic PDEs is quite deep. Set $(Y_s^{t,x})_{t \leqslant s \leqslant T}$ to be the solution of (3.24) with $X_t = x$. Then there exists a measurable function u

such that $u(s, X_s^{t,x}) = Y_s^{t,x}$. In particular, $u(t,x) = Y_t^{t,x}$ which is deterministic and continuous in both t and x. This u can be shown to be the unique viscosity solution of the QVI in Proposition 3. In particular, if the QVI admits a classical solution (for example if all the coefficients are $\mathcal{C}^3$ and bounded) then $u(t,x) \in \mathcal{C}^{1,2}([0,T] \times \mathbb{R}^d)$.

Remark 1. The quasi-variational inequality for optimal switching can be re-formulated as a coupled system of reflected BSDE's for $(Y^i)_{i \in \mathbb{Z}_M}$,

$$Y_t^i = \int_t^T \psi_i(s, X_s)\, ds + A_T^i - A_t^i - \int_t^T Z_s^i \cdot dW_s$$
$$Y_t^i \geqslant \max_{j \neq i} \{-C_{i,j} + Y_t^j\}.$$

However, the question of existence and uniqueness of solutions to such systems is difficult. In the special case of two regimes $M = 2$, Hamadène and Jeanblanc [44] explicitly proved existence by working with the difference process $Y^1 - Y^2$. For $M > 2$ it seems that the only available tools are either the analytic approach via viscosity solutions of the QVI or the recursive optimal stopping that we use.

Chapter 4

Numerical Solution

4.1 Dynamic Programming in Discrete Time

We now describe the numerical procedure for solving the optimal switching problem formulated recursively in (3.7). To be able to do numerical computation, we first pass from the continuous time to discrete time. Let $\{m\Delta t,\ m = 0, 1, \ldots, M^\sharp\}$, $\Delta t = \frac{T}{M^\sharp}$ be a discrete time grid. Switches are now allowed only at grid points, i.e. $\tau_k \in \mathcal{S}^\Delta = \{m\Delta t : 0 \leqslant m \leqslant M^\sharp\}$ and we label as $\mathcal{U}^\Delta$ the corresponding set of admissible strategies. This limiting of managerial flexibility is similar to looking at Bermudan options as approximation to American exercise rights. If the problem horizon is about $2-6$ months, a practical discretization is on the scale of $\Delta t = 3-12$ hours.

Let $t_1 = m\Delta t, t_2 = (m+1)\Delta t$ be two consecutive time steps. In discrete time, the Snell envelope is easily computed through the backward dynamic programming (DP) method. More precisely, the Snell envelope property of $J^k(t_1, x, i)$ reduces to

deciding between immediate switch at t_1 to some other regime j versus no switching and therefore waiting until t_2. Thus, (3.7) becomes

$$\begin{aligned} J^k(t_1, X_{t_1}, i) &= \max\Big(\mathbb{E}\big[\int_{t_1}^{t_2} \psi_i(s, X_s)\, ds + J^k(t_2, X_{t_2}, i)\big|\, \mathcal{F}_{t_1}\big],\, \mathcal{M}^{k,i}(t_1, X_{t_1})\Big) \\ &\simeq \max\Big(\psi_i(t_1, X_{t_1})\, \Delta t + \mathbb{E}\big[J^k(t_2, X_{t_2}, i)\big|\, \mathcal{F}_{t_1}\big], \\ &\qquad \max_{j\neq i}\big\{-C|i-j| + J^{k-1}(t_1, X_{t_1}, j)\big\}\Big). \end{aligned} \tag{4.1}$$

We see that to solve the problem it suffices to have a computationally efficient algorithm for evaluating the *conditional expectations* appearing in (4.1). Recall that the conditional expectation $\mathbb{E}[f(X_{t_2})|\, \mathcal{F}_{t_1}]$ is defined to be the $\mathcal{F}_{t_1}$-measurable random variable F which minimizes $\mathbb{E}[|f(X_{t_2}) - F|^2]$. On the other hand, if (X_t) is Markov, any such $\mathcal{F}_{t_1}$-measurable F may be written as $F = F(X_{t_1})$. Therefore, the conditional expectation can be viewed as simply a mapping $x \mapsto E_{t_1}(x) \triangleq \mathbb{E}[f(X_{t_2})|X_{t_1} = x] \in L^2(\mathbb{P})$. Hence, one possible numerical strategy for evaluating the former is to approximate the map E_{t_1} which is a well-studied statistical problem.

We concentrate on a particular variation first described by Longstaff and Schwartz [54] and Tsitsiklis and van Roy [67]. The idea is to *project* E_t onto a truncation of a basis of the Hilbert space $L^2(\mathbb{P})$. This finite projection is just a regression of E_t against the first N^B basis functions, i.e.

$$E_t(x) \simeq \hat{E}_t(x) = \sum_{j=1}^{N^B} \alpha_j B_j(x), \tag{4.2}$$

where $B_j(x)$ are the bases and α_j the $\mathbb{R}$-valued coefficients.

The computational advantage is that the projection can now be approximated with an empirical regression. The algorithm generates a large Monte Carlo sample $(x^\ell_{t_1}, x^\ell_{t_2})_{\ell=1}^{N^p}$ from the joint distribution of X_{t_1} and X_{t_2}. The empirical values $f(x^\ell_{t_2})$ are then regressed against $\{B_j(x^\ell_{t_1})\}$ to obtain the coefficients α_j after which we can use (4.2) to compute the conditional expectations.

We now extend this idea to allow for simultaneous recursive computation of all the J^k's. Begin by generating N^p sample paths x^ℓ of the discretized (X_t) process with a fixed initial condition $X_0 = x$. We will approximate $J^k(0, x, i)$ by $\frac{1}{N^p}\sum_\ell J^k(0, x_0^\ell, i)$. The pathwise values $J^k(t, x_t^\ell, i)$ are computed recursively in a backward fashion, starting with $J^k(T, x_T^\ell, i) = 0$. To implement (4.1), for a given step $t_1 = m\Delta t$ and regime i we regress the known future values $J^k((m+1)\Delta t, x_{(m+1)\Delta t}^\ell, i)$ onto the current underlying values $\{B_j(x_{m\Delta t}^\ell)\}$. As a result of regression we obtain a prediction $\hat{E}_{m\Delta t}\big[J^k(m\Delta t, \cdot, i)\big](x_{m\Delta t}^\ell)$ for the continuation value along the ℓ-th path. Comparing this with the current value $J^{k-1}(m\Delta t, x_{m\Delta t}^\ell, j)$ from a lower 'layer' $k-1$ for each switching choice j we can find the optimal decision at t_1. The computations are done bottom-up in k, so that indeed $J^{k-1}(m\Delta t, x_{m\Delta t}^\ell, j)$ is known when computing $J^k(m\Delta t, x_{m\Delta t}^\ell, i)$. The efficiency is maintained because we use *the same set* of paths to compute all the recursive conditional expectations. At a given layer k, the computations of J^k for different regimes i are independent of each other, and hence the errors only cumulate with respect to number of switches.

Recursion for τ^k

It turns out that for numerical efficiency rather than directly computing the value function, it is better to instead keep track of the optimal stopping time. This was the beautiful insight of Longstaff and Schwartz [54], as opposed to the Tsitsiklis and Van Roy paper [67] which implemented the just-described algorithm as is. Let $\tau^k(m\Delta t, X_{m\Delta t}, i)\cdot\Delta t$ correspond to the smallest optimal switching time for $J^k(m\Delta t, X_{m\Delta t}, i)$. In other words, the optimal future rewards are given by

$$J^k(m\Delta t, x, i) = \mathbb{E}\Big[\sum_{j=m}^{\tau^k} \psi_i(j\Delta t, X_{j\Delta t})\,\Delta t + \mathcal{M}^{k,i}(\tau^k\Delta t, X_{\tau^k\Delta t})\Big|\, X_{m\Delta t} = x\Big].$$

Then we have the analogue of (4.1) for τ^k:

$$\tau^k(m\Delta t, x_{m\Delta t}^\ell, i) = \begin{cases} \tau^k((m+1)\Delta t, x_{(m+1)\Delta t}^\ell, i), & \text{no switch;} \\ m, & \text{switch,} \end{cases} \tag{4.3}$$

and the set of paths on which we switch is given by $\{\ell\colon \hat{\jmath}^{\ell}(m\Delta t; i) \neq i\}$ with

$$\hat{\jmath}^{\ell}(t_1; i) = \arg\max_j \Big(-C_{i,j} + J^{k-1}(t_1, x^{\ell}_{t_1}, j),\ \psi_i(t_1, x^{\ell}_{t_1})\Delta t + \hat{E}_{t_1}\big[J^k(t_2, \cdot, i)\big](x^{\ell}_{t_1}) \Big). \tag{4.4}$$

The full recursive *pathwise* construction for J^k is

$$J^k(m\Delta t, x^{\ell}_{m\Delta t}, i) = \begin{cases} \psi_i(m\Delta t, x^{\ell}_{m\Delta t})\,\Delta t\ + J^k((m+1)\Delta t, x^{\ell}_{(m+1)\Delta t}, i), & \text{no switch;} \\ -C_{i,j} + J^{k-1}(m\Delta t, x^{\ell}_{m\Delta t}, j), & \text{switch to } j. \end{cases} \tag{4.5}$$

Observe that in this version the regression is used solely to update the optimal stopping times τ^k and the regressed values are never stored directly. This additional layer of random behavior helps to eliminate potential biases from the regression step.

Choosing the Basis Functions

The choice of appropriate basis functions (B_j) is rather heuristic. Several canonical choices have been proposed, including the Laguerre polynomials

$$B_j(x) = \mathrm{e}^{-x/2}\,\frac{\mathrm{e}^x}{j!}\frac{d^j(x^j e^{-x})}{dx^j}$$

from the original paper [54] and the indicator functions $B_j(x) = \mathbb{1}_{E_j}(x)$ of a partition of E [42]. A more automated approach was suggested by [46] based on neural networks. The idea is to use the logistic basis

$$B_j(x) = \frac{\mathrm{e}^{\alpha_j x}}{\mathrm{e}^{\alpha_1 x} + \ldots + \mathrm{e}^{\alpha_{N^B} x}},$$

where the coefficients (α_j) are picked using a factor analysis of the current data. In any case, the numerical precision can be greatly improved by customizing the basis. In particular, it helps to use basis functions that resemble the expected shape of the value function. In examples of this thesis ψ_i are usually linear and our favorite set of bases $B_j(x)$ are of the form x^p, $\mathrm{e}^{\alpha x}$ and $\max(x - K, 0)$.

In a multi-dimensional setting the simplest choice is to use tensor products of one-dimensional bases $\prod_k B_{j_k}$. This makes the required number of basis functions in principle exponential in the dimension of (X_t). However, if the rewards only depend on fixed linear combinations of components of X_t, even a linear number of bases might be sufficient to capture the relationship between the value function $J(t, x, i)$ and x. In practice, N^B as small as 5 or 6 normally suffices, and having more bases can often lead to worse numerical results due to overfitting.

4.1.1 Delay and Time Separation

An important feature of a realistic model is operational delay. Turning a physical plant on/off is not only costly, but also takes a significant amount of time. A typical plant requires $4 - 12$ hours of ramping up before it is operational. This delay is important if we look at markets with price spikes. Because there is a lag in taking a plant online, the operator is more risk averse, since he faces more uncertainty regarding the profitability of the spark spread by the time the plant is running. To approximate this phenomenon, we may want to assume that each switch takes a fixed amount of money $C_{i,j}$, as well as a fixed delay time δ, so that the plant is in its new regime only after $t + \delta$. Unfortunately, in continuous time, correct treatment of delay is cumbersome. The definition of the total reward obtained from strategy u must be changed from (2.2) to

$$H(x, i, [0, T]; u) = \int_0^T \psi_{u_{s-\delta}}(s, X_s)\, ds - C \int_0^T |du|_s$$

where we define $u_s = i$ for $-\delta \leqslant s \leqslant 0$. Note that u_s is still $\mathcal{F}_s$-measurable, but the corresponding payoff is only realized at $s + \delta$. The appearance of $\psi_{u_{s-\delta}}$ introduces acute technical difficulties, see for example work of Bar Ilan et al. [6] on irreversible real options with installation delay and Elsanosi et al. [37] on delayed optimal harvesting.

As an approximation of time delay we instead can impose *time separation* i.e. the constraint $\tau_k \geqslant \tau_{k-1} + \delta$[1] . This models the situation where the effect of the switch takes place immediately but then the plant is locked-up for a period of δ. Clearly, such situation is less risky than true delay because decisions have instantaneous impact, but it still limits the flexibility of the operator. With separation we also strictly speaking lose the Markov property of the value function. If we made a switch at t, time advances to $t + \delta$ while our hands are tied. Hence, (3.7) only holds at the switch times.

The advantage is that in discrete time, especially if δ is a multiple of Δt, separation is very convenient to implement. It means that instead of evaluating $\mathbb{E}[J^k(t + \Delta t, X_{t+\Delta t}, i)\big|\, \mathcal{F}_t]$ we now need to compute $\mathbb{E}[J^k(t + \delta, X_{t+\delta}, i)\big|\, \mathcal{F}_t]$ the conditional expectation $\delta/\Delta t$ steps ahead, but this is as easy as the original computation. Diagram 4.1 illustrates the dependency in space-time of the value functions with different delay settings.

Unlimited Number of Switches

If the physical problem has no upper limit on number of switches made, we can simplify the algorithm by replacing all J^k's with just J. In other words, we do not need to keep track of number of switches made. Indeed, on a discrete grid we can make at most M total switches, so that for $k > 2 \cdot M$, $J^k(m\Delta t, \cdot) \equiv J^{k-1}(m\Delta t, \cdot)$ for any m. This allows to flatten our construction of J^k in (4.1) and significantly speed up the computation.

Remark 2. In a real-life setting it is possible that the number of switches is limited a priori so that we really have to compute J^k for a fixed k. For instance, consider an operator who does not like switching due to political implications— he must not appear to be a price speculator. Thus, the management postulates that at most $\bar{K}$

[1]See also [17] for a related discussion in the case of energy swing options.

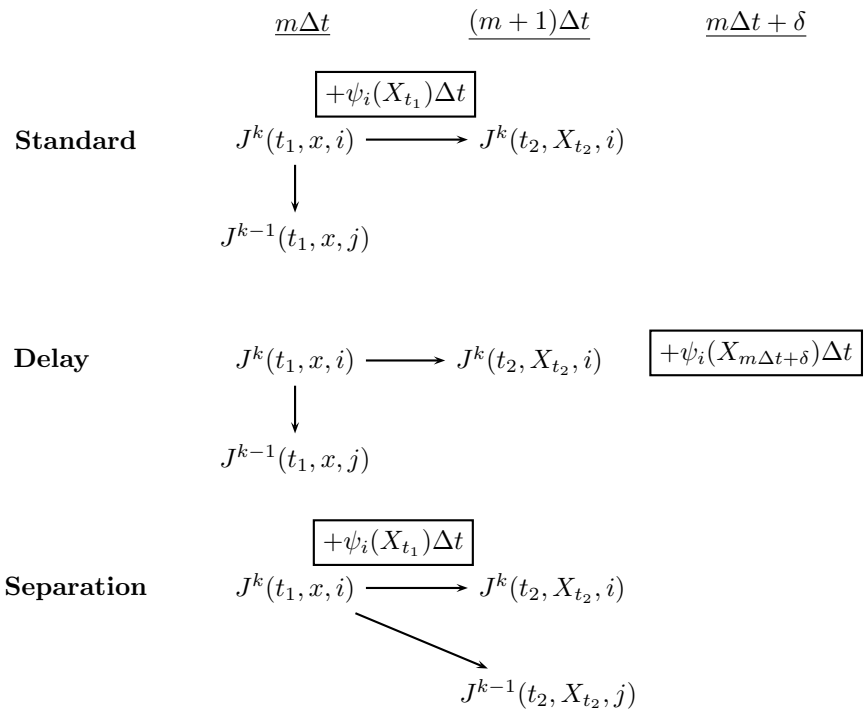


Figure 4.1: Decision Making with Different Delay Settings

switches can be made within a given month. If $\bar{K}$ is small, then $J^{\bar{K}} < J^{\infty}$ strictly and recursive computations are the *only* feasible approach.

4.1.2 Approximating the Switching Boundary

The Monte Carlo regression algorithm also produces an approximation to the switching boundary. Recall that the switching boundary specifies for each k, i the graph (t, x_t) such that the minimal optimal switching time corresponds to the first time of hitting this barrier. In our case, for each instant t and regime i we have $(M-1)$ boundaries containing the thresholds at which our optimal strategy at the current time step changes to regime j.

Since we keep track of optimal minimal τ, we can easily reconstruct the switching boundary by summarizing at the end of the algorithm the graph of $\tau^k(0, x_0, i)$ against X_t. Namely, the set

$$\{x^\ell_{m\Delta t} \colon \ell \text{ is such that } \tau^k(0, x^\ell, i) = m\} \tag{4.6}$$

defines the empirical region of switching from regime i at instant $m\Delta t$. The complement is the continuation set and we can determine the regime switched into by keeping track of $\hat{\jmath}^\ell$ from (4.4). The switching boundary can in turn be used in the forward direction to construct the optimal dispatch policy for any given path of (X_t). See Section 4.4 for a numerical example.

4.2 Summary of Algorithm

What follows is the summary of the numerical algorithm we use:

i. Select a set of basis functions (B_j) and algorithm parameters $\Delta t, M^\sharp, N^p, \bar{K}, \delta$.

ii. Generate N^p paths of the driving process: $\{x^\ell_{m\Delta t},\ m = 0, 1, \ldots, M^\sharp,\ \ \ell = 1, 2, \ldots, N^p\}$ with fixed initial condition $x^\ell_0 = x_0$.

iii. Initialize the value functions and minimal switching times $J^k(T, x^\ell_T, i) = 0$, $\tau^k(T, x^\ell_T, i) = M^\sharp\ \forall i, k$.

iv. Moving backward in time with $t = m\Delta t$, $m = M^\sharp, \ldots, 0$ repeat the following:

- Compute inductively the layers $k = 0, 1, \ldots, \bar{K}$ using (3.7). To evaluate the conditional expectation $\mathbb{E}\big[J^k(m\Delta t+\Delta t, \cdot, i)\big|\, \mathcal{F}_{m\Delta t}\big]$ regress $\{J^k(m\Delta t+\Delta t, x^\ell_{m\Delta t+\Delta t}, i)\}$ against current set of basis functions $\{B_j(x^\ell_{m\Delta t})\}_{j=1}^{N^B}$. Add the reward $\psi_i(m\Delta t, x^\ell_{m\Delta t}) \cdot \Delta t$ to the continuation value.
- Update the optimal switching times and value functions using (4.3), (4.4) and (4.5).

v. end Loop.

vi. Check back whether $\bar{K}$ switches are enough by comparing $J^{\bar{K}}$ and $J^{\bar{K}-1}$ (they should be equal).

vii. To obtain the free boundary, for each given $k, i, m\Delta t$ compute (4.6).

Observe that during the main loop we only need to store the buffer values $J(t,\cdot),\ldots,J(t+\delta,\cdot)$; and $\tau(t,\cdot),\ldots,\tau(t+\delta,\cdot)$.

We call the described algorithm the Longstaff-Schwartz Monte Carlo regression scheme, or just (LS). It is distinguished from the Tsitsiklis-van Roy (TvR) scheme which in step iv.2 directly uses the conditional expectations,

$$J^k(m\Delta t, x^\ell_{m\Delta t}, i) = \mathcal{M}^{k,i}(m\Delta t, x^\ell_{m\Delta t}) \\ \vee \Big[\hat{E}_{m\Delta t}\big[J^k((m+1)\Delta t,\cdot,i)\big](x^\ell_{(m+1)\Delta t}) + \psi_i(m\Delta t, x^\ell_{m\Delta t})\cdot\Delta t\Big]. \quad (4.7)$$

Algorithm Requirements

The speed of the algorithm is $O(M^2\cdot\bar{K}\cdot N^p\cdot M^\sharp)$ where M is the number of possible regimes, $\bar{K}$ maximum number of switches, N^p number of Monte Carlo paths, and $M^\sharp$ the number of discrete time stages used. The memory requirements are $O(N^p\cdot(M^\sharp + D\cdot\bar{K}\cdot M))$ where $D=\frac{\delta}{\Delta t}+1\geqslant 3$ is the buffer size and the two terms represent storage of sample paths and value functions respectively.

The algorithm complexity is quadratic in the number of regimes since we must check the possibility of switching from each mode i into each other mode j. The algorithm is linear in N^p since during the regression step we only deal with matrices of size $N^p\times N^B$. Because the other dimension is fixed, the number of arithmetic operations is linear in the bigger dimension.

In practice, the biggest constraint is in space, because the backward induction requires keeping the entire array $M^\sharp\times N^p$ of *sample paths* in memory. Even for reasonable values such as $M^\sharp=500, N^p=20,000$ this is already $40MB$ of storage.

4.3 Convergence

The presented algorithm for computing J^k has several layers of approximations. We now take a systematic tally of all the errors involved. Looking back four types of errors can be identified: error due to discretizing the SDE, error due to restricting switching times, projection error and Monte Carlo sampling error. We shall address each of these errors in turn and present the current state-of-the-art regarding error bounds in terms of Δt and the number of paths N^p.

The framework of backward stochastic differential equations from Section 3.7 has been the most successful for analyzing convergence properties of optimal stopping algorithms. The basic TvR scheme (4.7) is essentially a simple algorithm for solving discrete-time reflected BSDE's, a topic that has been a very active area of research, see e.g. [11, 21, 42]. The expected convergence properties of the TvR scheme are not as good as of the LS scheme, however the former is much more amenable to analysis.

The BSDE approach starts from (3.24) and constructs a discrete time approximation to the pair (Y, Z). As we will see, this again reduces to computing conditional expectations. The initial analysis of Bouchard and Touzi [11] used independent approximations for each conditional expectation making errors at each time step uncorrelated. More recently, Gobet et al. [42] extended this analysis to the LS/TvR approach of using a single set of (X_t)-paths to approximate all the conditional expectations. To simplify our proofs let us restate our framework with slightly stronger assumptions. Recall that in terms of BSDE's the problem reduces to finding $Y_t^{k,i} \equiv J^k(t, X_t^x, i)$ (again we suppress dependence on x) where

$$X_t = x + \int_0^t \mu(X_s)\, ds + \int_0^t \sigma(X_s) \cdot dW_s, \tag{4.8}$$

$$Y_t^{k,i} = \int_t^T \psi_i(s, X_s)\, ds + A_T - A_t - \int_t^T Z_s \cdot dW_s,$$

$$Y_t^{k,i} \geqslant \mathcal{M}_t^{k,i} = \max_{j \neq i} \big\{ -C_{i,j} + Y_t^{k-1,j} \big\}, \quad \text{and} \quad \int_0^T \big(Y_t^{k,i} - \mathcal{M}_t^{k,i} \big)\, dA_t = 0, \tag{4.9}$$

As usual, i represents the initial regime and k counts the number of switches, so when $k = 0$ we simply have no barrier in (4.9). Because of the Markovian setting $Y_t^{k,i}$ (and hence also $\mathcal{M}_t^{k,i}$) are functions of X_t^x, a fact that allows to connect the backward and forward components.

Assumption 5. *$\mu(x)$ and $\sigma(x)$ are twice continuously differentiable, bounded and Lipschitz. The payoffs $\psi_i(x)$ are time-homogeneous, continuously differentiable and Lipschitz.*

The convergence proof will proceed in several steps. In Step 1, we will estimate the error resulting from time discretization— discretizing (X_t) using Euler scheme and $(Y_t^{k,i})$ using the standard backward scheme. Simultaneously we restrict the switching times to occur only at the discrete time grid points. We show that the error from this procedure which produces $(X^\Delta, Y^{k,i,\Delta}, Z^{k,i,\Delta})$ is $O(\sqrt{\Delta t})$. In Step 2, we replace the conditional expectations with a projection $\mathcal{P}_m$ on a finite dimensional orthonormal family in $L^2(\mathbb{P})$. Hence, $Y^{k,i,\Delta}$ is approximated by $\hat{Y}^{k,i}$. The resulting error is $O(\Delta t^{-k} \cdot (\text{mean projection error}))$. Finally, in Step 3 we further approximate the projections by an empirical regression using N realizations of the paths $(x^\ell_{m\Delta t}$, $\ell = 1, \ldots, N)$. We conjecture that the error here is $O((\Delta t \cdot N)^{-1/2})$, the expected rate for Monte Carlo methods.

This section will be heavy on notation, and we will often drop unnecessary subscripts and superscripts to prevent clutter. Throughout M and N are used to denote the number of discrete time periods and total number of Monte Carlo simulations respectively. The time steps are indexed by m and the paths by ℓ. In the proofs below C is a generic constant that may change from line to line, and t_1 and t_2 are two generic adjacent time steps, $t_1 = m\Delta t, t_2 = t_1 + \Delta t$. Finally, we use the standard notation $X \vee Y \triangleq \max(X, Y)$ and $\|X\|_p = (\mathbb{E}[|X|^p])^{1/p}, p \geqslant 1$ for the L^p-norm of the random variable X. At this point let us also recall the following

Lemma 3. *If positive sequences* $(a_m), (b_m), (c_m)$, $0 \leqslant m \leqslant M$ *satisfy* $a_{m-1} \leqslant \max\{(1 + C\Delta t)a_m + b_{m-1},\, c_{m-1}\}$ *then*

$$a_m \leqslant \mathrm{e}^{C(T-m\Delta t)} \left[a_M + \sum_{j \geqslant m} b_j + \max_{j \geqslant m} c_j \right]. \tag{4.10}$$

This is just a straightforward extension of the discrete version of Gronwall's inequality for the case where term-by-term maximum is taken. This lemma will be key for L^2 estimates when combined with Young's inequality: $(a+b)^2 \leqslant (1+\gamma)a^2 + (1+\frac{1}{\gamma})b^2$ for any $\gamma > 0$.

4.3.1 Discretization Error

To discretize, we first use the classical Euler scheme for the process (X_t)[2] :

$$\begin{cases} X^{\Delta}_{t_2} &= X^{\Delta}_{t_1} + \mu(X^{\Delta}_{t_1})\Delta t + \sigma(X^{\Delta}_{t_1}) \cdot (W_{t_2} - W_{t_1}), \\ X^{\Delta}_{t} &= X^{\Delta}_{t_1} + \mu(X^{\Delta}_{t_1})(t - t_1) + \sigma(X^{\Delta}_{t_1}) \cdot (W_{t} - W_{t_1}), \quad \text{for } t \in (t_1, t_2). \end{cases} \tag{4.11}$$

Then a standard L^p-bound is [11, Lemma 3.2]

$$\limsup_{\Delta t \to 0} \frac{1}{\sqrt{\Delta t}} \left\| \sup_{0 \leqslant t \leqslant T} |X_t - X^{\Delta}_t| \right\|_p < \infty \quad \forall p \geqslant 1. \tag{4.12}$$

Assuming (4.11) for (X^{Δ}_t), the discretized versions of $(Y^{k,i}, Z^{k,i})$ from (4.8), which we label $(Y^{k,i,\Delta}, Z^{k,i,\Delta})$, and in an obvious way $\mathcal{M}^{k,i,\Delta}$ solve

$$Y^{k,i,\Delta}_{t_1} = \max\Big\{ \mathcal{M}^{k,i,\Delta}_{t_1},\, \mathbb{E}\big[Y^{k,i,\Delta}_{t_2} \big|\, \mathcal{F}_{t_1} \big] + \psi_i(X^{\Delta}_{t_1})\Delta t \Big\}, \quad Y^{k,i,\Delta}_{T} = 0, \tag{4.13}$$

$$Z^{k,i,\Delta}_{t_1} = \mathbb{E}\Big[Y^{k,i,\Delta}_{t_2} \cdot \frac{(W_{t_2} - W_{t_1})}{\Delta t} \Big|\, \mathcal{F}_{t_1} \Big]. \tag{4.14}$$

The alternative

$$Y^{k,i\Delta}_{t_1} = \max\Big\{ \mathcal{M}^{k,i,\Delta}_{t_1},\, \mathbb{E}\Big[Y^{k,i,\Delta}_{t_2} + \psi_i(X^{\Delta}_{t_2})\Delta t \big|\, \mathcal{F}_{t_1} \Big] \Big\}$$

has also been used by some authors, e.g. Bally et al. [4]. There is not much difference for our setting, however (4.13) is better for the general situation where the generator

[2] If (X_t) is Gaussian this is not needed, but we cover the general setting since the backward components must be discretized in any case.

ψ_i may depend on $Y^{k,i}$. From a financial point of view (4.13) means that the payoff between today and tomorrow is certain. It is not clear which assumption is more realistic but numerically for small Δt the differences seem to be negligible.

The next theorem is an analogue of Theorem 7.1 in Bouchard and Touzi [11]. It shows that the total discretization error is $O(\sqrt{\Delta t})$. We provide a full proof for completeness.

Theorem 6. *For all $p \geqslant 1$,*

$$\limsup_{\Delta t \to 0} \frac{1}{\sqrt{\Delta t}} \sup_{0 \leqslant m \leqslant M} \left\| Y^{k,i,\Delta}_{m\Delta t} - Y^{k,i}_{m\Delta t} \right\|_p < \infty.$$

Proof. As in all subsequent proofs we will proceed by induction on k remembering that the first level has no barrier and reduces to the standard BSDE situation. Let $\mathcal{S}^\Delta(m\Delta t) = \{j\Delta t \colon m \leqslant j \leqslant M\}$ be the set of our discretized stopping times after $m\Delta t$ and define the auxiliary processes

$$R^{k,i}_{m\Delta t} \triangleq \operatorname*{ess\,sup}_{\tau \in \mathcal{S}^\Delta(m\Delta t)} \mathbb{E}\Big[\mathcal{M}^{k,i}_\tau + \Delta t \sum_{j=m}^{M} \mathbb{1}_{\tau > j\Delta t} \cdot \psi_i(X_{j\Delta t}) \Big| \mathcal{F}_{m\Delta t} \Big],$$

$$L^{k,i}_{m\Delta t} \triangleq \operatorname*{ess\,sup}_{\tau \in \mathcal{S}^\Delta(m\Delta t)} \mathbb{E}\Big[\mathcal{M}^{k,i}_\tau + \int_{m\Delta t}^{\tau} \psi_i(X_s)\, ds \Big| \mathcal{F}_{m\Delta t} \Big].$$

Hence, $L^{k,i}$ discretizes the stopping rule, while $R^{k,i}$ converts the integrated reward into a sum. We will use

$$|Y^{k,i,\Delta}_t - Y^{k,i}_t| \leqslant |Y^{k,i,\Delta}_t - R^{k,i}_t| + |R^{k,i}_t - L^{k,i}_t| + |L^{k,i}_t - Y^{k,i}_t|$$

and separately estimate the three terms of the right hand side. First, for any $\tau \in \mathcal{S}(m\Delta t)$, set $\tilde{\tau} = \lceil \frac{\tau}{\Delta t} \rceil \cdot \Delta t \in \mathcal{S}^{\Delta}(m\Delta t)$ so that $\tilde{\tau}$ rounds up τ to the next time point on the grid. Then using the fact that ψ_i's are Lipschitz we have

$$\begin{aligned}
|Y^{k,i}_{m\Delta t} - L^{k,i}_{m\Delta t}| &\leqslant \operatorname*{ess\,sup}_{\tau \in \mathcal{S}(m\Delta t)} \mathbb{E}_m \left[|\mathcal{M}^{k,i}_{\tau} - \mathcal{M}^{k,i}_{\tilde{\tau}}| + \int_{\tau}^{\tilde{\tau}} |\psi_i(X_s)|\, ds \right] \\
&\leqslant \operatorname*{ess\,sup}_{\tau \in \mathcal{S}(m\Delta t)} \max_j \mathbb{E}_m \Big[2 \cdot \int_{\tau}^{\tilde{\tau}} |\psi_j(X_s)|\, ds \Big] \\
&\leqslant 2\Delta t \cdot \mathbb{E}_m \Big[\sup_{m\Delta t \leqslant s \leqslant T} \max_j |\psi_j(X_s)| \Big] \\
\implies \quad \|Y^{k,i}_{m\Delta t} - L^{k,i}_{m\Delta t}\|_p &\leqslant C\Delta t \, \big\| \sup_{0 \leqslant s \leqslant T} |X_s| \big\|_p \leqslant C\Delta t.
\end{aligned}$$

Next under Assumption 5 by a standard argument $\max_{0 \leqslant m \leqslant M} \|R^{k,i}_{m\Delta t} - L^{k,i}_{m\Delta t}\|_p \leqslant C\Delta t$, see e.g. Lemma 4 in [4]. Observe that both estimates above are uniform in k because the approximation concerns $\psi_i(X_t)$ only.

The bulk of the error is hidden inside the difference of $R^{k,i}_{m\Delta t}$ and $Y^{k,i,\Delta}_{m\Delta t}$ which use different barriers and different X's. Applying $|(X_1 \vee Y_1) - (X_2 \vee Y_2)| \leqslant |X_1 - X_2| \vee |Y_1 - Y_2|$ on $|\mathcal{M}^{k,i}_{\tau} - \mathcal{M}^{k,i,\Delta}_{\tau}| = |\max_{j \neq i}(Y^{k-1,j}_{\tau} - C_{i,j}) - \max_{j \neq i}(Y^{k-1,j,\Delta}_{\tau} - C_{i,j})|$ we have

$$\begin{aligned}
|R^{k,i}_{m\Delta t} - Y^{k,i,\Delta}_{m\Delta t}| &\leqslant \operatorname*{ess\,sup}_{\tau \in \mathcal{S}^{\Delta}(m\Delta t)} \mathbb{E}_m \Big[|\mathcal{M}^{k,i}_{\tau} - \mathcal{M}^{k,i,\Delta}_{\tau}| + \Delta t \sum_{j=m}^{M} \mathbb{1}_{\tau > j\Delta t} \cdot |\psi_i(X^{\Delta}_{j\Delta t}) - \psi_i(X_{j\Delta t})| \Big] \\
&\leqslant \mathbb{E}_m \Big[\max_{m \leqslant j \leqslant M} \max_{j' \neq i} |Y^{k-1,j'}_{j\Delta t} - Y^{k-1,j',\Delta}_{j\Delta t}| + C|X^{\Delta}_{j\Delta t} - X_{j\Delta t}| \Big] \quad \text{or} \\
\|R^{k,i}_{m\Delta t} - Y^{k,i,\Delta}_{m\Delta t}\|_p &\leqslant C\sqrt{\Delta t} + \big\| \max_{0 \leqslant j \leqslant M} \max_{j' \neq i} |Y^{k-1,j'}_{j\Delta t} - Y^{k-1,j',\Delta}_{j\Delta t}| \big\|_p \qquad \text{by (4.12).}
\end{aligned}$$

Combining the three individual bounds we finally obtain,

$$\big\| \max_{0 \leqslant m \leqslant M} |Y^{k,i,\Delta}_{m\Delta t} - Y^{k,i}_{m\Delta t}| \big\|_p \leqslant C\sqrt{\Delta t} + \big\| \max_{0 \leqslant m \leqslant M} \max_{j' \neq i} |Y^{k-1,j'}_{m\Delta t} - Y^{k-1,j',\Delta}_{m\Delta t}| \big\|_p$$

where the constant C is intrinsic to (X_t) and ψ_i's. By induction on k, the $O(\sqrt{\Delta t})$ error propagates through for any fixed level k. □

The above error due to replacing the 'American' switching policy with a 'Bermudan' one has also been studied by Dupuis and Wang [29]. Using Brownian local time

techniques they directly show that for a one-dimensional optimal stopping problem the discretized value functions converge with rate $O(\Delta t)$ and give precise breakdown of the error sources. They furthermore demonstrate that the optimal stopping rules converge only with rate $O(\sqrt{\Delta t})$, a result that is unavailable with the BSDE formulation.

4.3.2 Projection Error

For the next step we approximate the conditional expectation $\mathbb{E}_{t_1} \triangleq \mathbb{E}[\cdot|\,\mathcal{F}_{t_1}]$ by a finite-dimensional projection $\mathcal{P}_{t_1}$ onto the set of bases $B_{t_1} = B_{t_1}(X_{t_1})$ (the latter is a vector of length N^B, but we do not write out the individual components). We label by α_{t_1} the resulting coefficients which form a random vector in $\mathbb{R}^{N^B}$, and by $\hat{Y}_{t_1}^{k,i}$ the approximation of $Y_{t_1}^{k,i,\Delta}$. Hence the equation (4.13) is replaced with

$$\begin{aligned}\hat{Y}_{t_1}^{k,i} &= \Big(\mathcal{P}_{t_1}[\hat{Y}_{t_2}^{k,i}] + \psi_i(X_{t_1})\Delta t\Big) \vee \hat{\mathcal{M}}_{t_1}^{k,i} \\ &= \Big(\alpha_{t_1}\cdot B_{t_1} + \psi_i(X_{t_1})\,\Delta t\Big) \vee \hat{\mathcal{M}}_{t_1}^{k,i}\end{aligned} \tag{4.15}$$

where α_{t_1} is given by $\alpha_{t_1} = \arg\min_\alpha \mathbb{E}\big[|\hat{Y}_{t_2}^{k,i} - \alpha\cdot B_{t_1}|^2\big]$.

We first check that for a fixed k, $\hat{Y}_{m\Delta t}^{k,i}$ is uniformly L^2-integrable. Using

$$\mathbb{E}|(X_1\vee Y_1) - (X_2\vee Y_2)|^2 \leqslant (1+\gamma)\mathbb{E}|X_1 - X_2|^2 \vee (1+\frac{1}{\gamma})\mathbb{E}|Y_1 - Y_2|^2 \tag{4.16}$$

as well as the fact that $\mathcal{P}_{t_1}$ is an L^2-contraction and Young's inequality we obtain

$$\begin{aligned}\mathbb{E}[|\hat{Y}_{t_1}^{k,i}|^2] &= \mathbb{E}\left[|\left(\mathcal{P}_{t_1}[\hat{Y}_{t_2}^{k,i}] + \Delta t\,\psi_i(X_{t_1})\right)\vee \hat{\mathcal{M}}_{t_1}^{k,i}|^2\right] \\ &\leqslant \Big((1+C\Delta t)\,\mathbb{E}\big[\mathcal{P}_{t_1}[\hat{Y}_{t_2}^{k,i}]^2\big] + (\Delta t^2 + \frac{\Delta t}{C})\,\mathbb{E}\left[\psi_i(X_{t_1})^2\right]\Big) \vee (1+\frac{1}{C\Delta t})\mathbb{E}|\hat{\mathcal{M}}_{t_1}^{k,i}|^2 \\ &\leqslant \Big((1+C\Delta t)\,\mathbb{E}\big[|\hat{Y}_{t_2}^{k,i}|^2\big] + C\Delta t(1+\|x\|^2)\Big) \vee C(1+\frac{1}{\Delta t})\max_{j\neq i}\mathbb{E}[|\hat{Y}_{t_1}^{k-1,j}|^2]\end{aligned}$$

so that by Gronwall's lemma

$$\max_{0\leqslant m\leqslant M}\mathbb{E}[|\hat{Y}_{m\Delta t}^{k,i}|^2] \leqslant C\big(1+\|x\|^2 + \frac{1}{\Delta t}\max_{j\neq i}\max_{0\leqslant m\leqslant M}\mathbb{E}[|\hat{Y}_{m\Delta t}^{k-1,j}|^2]\big). \tag{4.17}$$

Completing the induction on k, $\max_{0\leqslant m\leqslant M}\mathbb{E}[|\hat{Y}^{k,i}_{m\Delta t}|^2]\leqslant C(\Delta t)^{-k}(1+\|x\|^2)$. Even if this bound deteriorates as k increases, it still provides useful a priori estimates on the regression result $\check{Y}_{t_1}\triangleq\alpha_{t_1}\cdot B_{t_1}$. By (4.17) and orthonormality of B_{t_1} we have

$$\mathbb{E}|\check{Y}^{k,i}_{t_1}|^2=\mathbb{E}[|\alpha_{t_1}|^2|B_{t_1}|^2]=|\alpha_{t_1}|^2\leqslant C(1+\|x\|^2)$$

and consequently,

$$|\check{Y}^{k,i}_{t_1}|\leqslant|\alpha_{t_1}|\cdot|B_{t_1}(X_{t_1})|\leqslant|B_{t_1}(X_{t_1})|\sqrt{C(1+\|x\|^2)}. \tag{4.18}$$

Thus, we can construct a truncation function T_{t_1} such that $|\check{Y}^{k,i}_{t_1}|\leqslant T_{t_1}(X_{t_1})$ and $\mathbb{E}[T_{t_1}(X_{t_1})^2]<\infty$. This truncation is used to bound back α_{t_1} via

$$|\alpha_{t_1}|^2=\mathbb{E}|\check{Y}^{k,i}_{t_1}|^2\leqslant\mathbb{E}[T_{t_1}(X_{t_1})^2].$$

We now investigate in more detail the build-up of regression errors. Let $\mathcal{R}_{t_1}$ denote the remainder after the projection $\mathcal{P}_{t_1}$ so that for any $X\in\mathcal{F}_{t_1}$, $X=\mathcal{P}_{t_1}(X)+\mathcal{R}_{t_1}(X)$, and the two latter terms are orthogonal in L^2. The following theorem is similar to Theorem 2 in Gobet et al. [42].

Theorem 7. *With the notation of* (4.15)*, and* $Y^{k,i}\equiv Y^{k,i,\Delta}$*,*

$$\max_{0\leqslant m\leqslant M}\mathbb{E}\big[|\hat{Y}^{k,i}_{m\Delta t}-Y^{k,i}_{m\Delta t}|^2\big]\leqslant C\sum_{m=0}^{M}\mathbb{E}\big[\mathcal{R}_j(\mathbb{E}_m[Y^{k,i}_{(m+1)\Delta t}])^2\big]+\frac{C}{\Delta t}\max_{0\leqslant m\leqslant M}\max_{j\neq i}\mathbb{E}\big[|\hat{Y}^{k-1,j}_{m\Delta t}-Y^{k-1,j}_{m\Delta t}|^2\big].$$

Proof. The theorem shows that the regression errors add up when moving across a fixed level k and get multiplied by $\frac{1}{\Delta t}$ when moving down to level $k-1$. Set $\eta^{k,i}_{t_1}\triangleq\mathbb{E}\left[|\hat{Y}^{k,i}_{t_1}-Y^{k,i}_{t_1}|^2\right]$. We will show that

$$\eta^{k,i}_{t_1}\leqslant(1+C\Delta t)\big(\eta^{k,i}_{t_2}+\mathbb{E}[\mathcal{R}_{t_1}(\mathbb{E}_{t_1}[Y^{k,i}_{t_2}])^2]\big)\vee(1+\frac{1}{C\Delta t})\max_{j\neq i}\eta^{k-1,j}_{t_1}.$$

Combined once again with Lemma 3 this is enough to prove the theorem. To show the claimed inequality, re-write

$$\eta_{t_1}^{k,i} = \mathbb{E}\Big|\Big\{\mathcal{P}_{t_1}(\hat{Y}_{t_2}^{k,i}) + \psi_i(t_1, X_{t_1})\Delta t \vee \hat{\mathcal{M}}_{t_1}^{k,i}\Big\} - \Big\{\mathbb{E}_{t_1}(Y_{t_2}^{k,i}) + \psi_i(t_1, X_{t_1})\Delta t \vee \mathcal{M}_{t_1}^{k,i}\Big\}\Big|^2$$
$$\leqslant \mathbb{E}\Big[|\mathcal{P}_{t_1}(\hat{Y}_{t_2}^{k,i}) - \mathbb{E}_{t_1}(Y_{t_2}^{k,i})|^2 \vee |\hat{\mathcal{M}}_{t_1}^{k,i} - \mathcal{M}_{t_1}^{k,i}|^2\Big] \tag{4.19}$$

Since $\mathcal{P}_{t_1}$ is an L^2-projection, for any Z, $\mathcal{P}_{t_1}(Z) = \mathcal{P}_{t_1}(\mathbb{E}_{t_1}[Z])$ and so $\mathbb{E}_{t_1}(Y_{t_2}^{k,i}) = \mathcal{P}_{t_1}(Y_{t_2}^{k,i}) + \mathcal{R}_{t_1}(\mathbb{E}_{t_1}[Y_{t_2}^{k,i}])$ implying

$$\mathbb{E}|\mathcal{P}_{t_1}(\hat{Y}_{t_2}^{k,i}) - \mathbb{E}_{t_1}[Y_{t_2}^{k,i}]|^2 \leqslant \mathbb{E}[\mathcal{R}_{t_1}(\mathbb{E}_{t_1}[Y_{t_2}^{k,i}])^2] + \mathbb{E}\left[|\hat{Y}_{t_2}^{k,i} - Y_{t_2}^{k,i}|^2\right]$$

by the orthogonality of the remainder $\mathcal{R}_{t_1}$. Substituting into (4.19) and using (4.16)

$$\eta_{t_1}^{k,i} \leqslant (1 + C\Delta t)\big(\mathbb{E}[\mathcal{R}_{t_1}(\mathbb{E}_{t_1}[Y_{t_2}^{k,i}])^2] + \eta_{t_2}^{k,i}\big) \vee (1 + \frac{1}{C\Delta t})\mathbb{E}|\hat{\mathcal{M}}_{t_1}^{k,i} - \mathcal{M}_{t_1}^{k,i}|^2.$$

as desired. □

In a more general way we can think of $\mathcal{P}_{t_1}$ as any approximation (not necessarily a projection) of the conditional expectation. For example, $\mathcal{P}_{t_1}$ may be an empirical Monte Carlo average. In this case the above theorem is slightly modified because we can no longer use the orthogonality of the projection errors. We state the following result which originally appeared in Bouchard and Touzi [11]:

Theorem 8. *For any m, k, i,*

$$\|\hat{Y}_{m\Delta t}^{k,i} - Y_{m\Delta t}^{k,i}\|_p \leqslant \frac{C}{\Delta t}\max_{m\leqslant j\leqslant M}\Big(\big\|(\mathbb{E}_j - \mathcal{P}_j)(\hat{Y}_{(j+1)\Delta t}^{k,i})\big\|_p + \max_{j'\neq i}\|\hat{Y}_{j\Delta t}^{k-1,j'} - Y_{j\Delta t}^{k-1,j'}\|_p\Big).$$

Proof. Using same techniques as for (4.19),

$$|Y_{t_1}^{k,i} - \hat{Y}_{t_1}^{k,i}| \leqslant \Big(|(\mathbb{E}_{t_1} - \mathcal{P}_{t_1})(\hat{Y}_{t_2}^{k,i})| + \mathbb{E}_{t_1}|Y_{t_2}^{k,i} - \hat{Y}_{t_2}^{k,i}|\Big) \vee \max_{j\neq i}|Y_{t_1}^{k-1,j} - \hat{Y}_{t_1}^{k-1,j}|$$

$$\|Y_{t_1}^{k,i} - \hat{Y}_{t_1}^{k,i}\|_p \leqslant (1 + C\Delta t)\Big(\|(\mathbb{E}_{t_1} - \mathcal{P}_{t_1})(\hat{Y}_{t_2}^{k,i})\|_p + \|Y_{t_2}^{k,i} - \hat{Y}_{t_2}^{k,i}\|_p\Big) \vee (1 + \frac{1}{C\Delta t})\max_{j'\neq i}\|Y_{t_1}^{k-1,j'} - \hat{Y}_{t_1}^{k-1,j'}\|_p$$

which by Lemma 3 implies the result. □

4.3.3 Sampling Error

In the final step we simulate N sample paths $\{x^\ell_{m\Delta t}, \ell = 1, \dots, N\}$ and replace the projection $\mathcal{P}_m$ by an empirical regression. We call the resulting regression coefficients $\alpha^N_{t_1}$ and the resulting basis functions $B^\ell_{t_1} = B_{t_1}(x^\ell_{t_1})$. The approximation of $\hat{Y}^{k,i}$ along the ℓ-th path is labeled as $\hat{Y}^{k,i,\ell}$ and satisfies

$$\hat{Y}^{k,i,\ell}_{t_1} = \Big(T^\ell_{t_1}(\alpha^N_{t_1} \cdot B^\ell_{t_1}) + \psi_i(x^\ell_{t_1})\Delta t\Big) \vee \hat{\mathcal{M}}^{k,i,\ell}_{t_1}, \tag{4.20}$$

$$\alpha^N_{t_1} = \arg\min_\alpha \frac{1}{N}\sum_{\ell=1}^N (\hat{Y}^{k,i,\ell}_{t_2} - \alpha \cdot B^\ell_{t_1})^2. \tag{4.21}$$

The empirical truncation $T^\ell_{t_1}(x)$ is an analytical device relying on the a priori estimate (4.18) to make sure that the simulation Y's are bounded. Letting $\check{Y}^{k,i,\ell}_{t_1} = \alpha^N_{t_1} \cdot B^\ell_{t_1}$, it is taken such that $T^\ell_{t_1}(\check{Y}^{k,i}_{t_1}) = T^\ell_{t_1}(\alpha_{t_1} \cdot B_{t_1}(x^\ell_{t_1})) = \alpha_{t_1} \cdot B_{t_1}(x^\ell_{t_1})$ and $T^\ell_{t_1}(\alpha^N_{t_1} \cdot B_{t_1}(x^\ell_{t_1})) \leqslant 2 \cdot T_{t_1}(x^\ell_{t_1})$. In other terms, $|\check{Y}^{k,i,\ell}| \leqslant 2 \cdot |\check{Y}^{k,i}|$. Note that T^ℓ is random since it depends on α_{t_1}.

Full analysis of the Monte Carlo sampling error $\mathbb{E}|\hat{Y}^{k,i}_{t_1} - hY^{k,i,N}_{t_1}|$ where $\hat{Y}^{k,i,N}_{t_1} = \frac{1}{N}\sum_{\ell=1}^N \hat{Y}^{k,i,\ell}_{t_1}$ appears to be intractable. The major difficulty is the presence of the recursively defined $\hat{\mathcal{M}}^{k,i,\ell}$ which causes the sampling error to propagate in a nonlinear fashion. Moreover, it is not clear how to compare the empirical $\hat{Y}^{k,i,N}$ which is a sum of max-terms in (4.20) with $\hat{Y}^{k,i}$ which is a single max-term. Looking closely the sampling error has three components that are closely intertwined— error due to using $\alpha^N_{t_1}$ rather than α_{t_1}, error due to using $\hat{\mathcal{M}}^{k,i,\ell}$ rather than $\hat{\mathcal{M}}^{k,i}$, and error due to potentially choosing the wrong side in the max-comparison, i.e. picking $T^\ell_{t_1}(\check{Y}^{k,i,\ell}_{t_1}) + \psi_i(x^\ell_{t_1})\Delta t$ over $\hat{\mathcal{M}}^{k,i,\ell}$, when $\alpha_{t_1} \cdot B^\ell_{t_1} + \psi_i(x^\ell_{t_1})\Delta t < \hat{\mathcal{M}}^{k,i,\ell}$.

We do not know how to isolate these error sources, so let us summarize the results of other papers that deal with simpler cases. The closest analogue is the paper of Gobet et al. [42] who study the non-reflected case. Their main idea is to denote by

$B^N_{t_1} = \frac{1}{N}\sum_{\ell=1}^N B^\ell_{t_1}(B^\ell_{t_1})^T \in \mathbb{R}^{N^B\times N^B}$ the sample regression matrix and to define the 'good' event

$$\mathbf{A} \triangleq \{\forall m\colon \|B^N_{m\Delta t} - Id\| \leqslant \Delta t\}.$$

Remembering that the basis is orthonormal, for any fixed Δt, $\mathbb{P}(\mathbf{A}) \to 1$ as $N \to \infty$ by the law of large numbers. On the other hand, using the truncation T_{t_1} and the fact that the regression matrix is close to identity on $\mathbf{A}$, we have (we abuse notation and use $\hat{Y}^{k,i,N}$ even after dropping the reflection terms involving $\mathcal{M}$ in (4.20) and (4.15))

$$\begin{aligned}
\mathbb{E}|\hat{Y}^{k,i,N}_{t_1} - \hat{Y}^{k,i}_{t_1}|^2 &\leqslant \mathbb{E}[\mathbb{1}_{\mathbf{A}^C}|\hat{Y}^{k,i,N}_{t_1} - \hat{Y}^{k,i}_{t_1}|^2] + \mathbb{E}\big[\mathbb{1}_{\mathbf{A}}|\hat{Y}^{k,i,N}_{t_1} - \hat{Y}^{k,i}_{t_1}|^2\big] \\
&\leqslant \mathbb{E}[\mathbb{1}_{\mathbf{A}^C}|2\cdot T_{t_1}(X_{t_1}) + T_{t_1}(X_{t_1})|^2] + \mathbb{E}[\mathbb{1}_{\mathbf{A}}\|B^N_{t_1}\|^2|\alpha^N_{t_1} - \alpha_{t_1}|^2] \\
&\leqslant 9\cdot\mathbb{P}[\mathbf{A}^c]^{1/2}\cdot\mathbb{E}[T_{t_1}(X_{t_1})^4]^{1/2} + (1+\Delta t)\mathbb{E}[\mathbb{1}_{\mathbf{A}}|\alpha^N_{t_1} - \alpha_{t_1}|^2].
\end{aligned}$$

Hence on $\mathbf{A}$ it suffices to estimate the error from the wrong regression coefficients $\alpha^N_{t_1}$. Using Gronwall-type inequality Gobet et al. [42, Theorem 3] then show that

$$\max_{0\leqslant m\leqslant M}\mathbb{E}\Big[\mathbb{1}_{\mathbf{A}}|\alpha^N_{m\Delta t} - \alpha_{m\Delta t}|^2\Big] \leqslant \frac{C}{\Delta t\cdot N}$$

implying that the rate of convergence is $O(N^{-1/2})$ on $\mathbf{A}$. It still remains unclear how fast does the 'good' event itself grow; observe that $\mathbf{A}$ requires the regression matrix norm to be small uniformly in $m\Delta t$.

From a different angle, Bouchard and Touzi [11, Theorem 4.1] show that if one directly approximates $\mathbb{E}_{t_1}$ with a Monte Carlo simulation then the resulting rate of convergence in L^p-norm is $O(\Delta t^{-1-d/4p}\cdot N^{-2p})$ both for the non-reflected case and a single-level reflection. The technique is again to use a priori bounds on the empirical Y's and Gronwall's inequality.

On a more abstract level, Clément et al. [22] have shown that fixing N^B and using the Longstaff-Schwartz scheme (4.3) of approximating the optimal switching time, the asymptotic sampling error as $N \to \infty$ has mean zero and a Gaussian

distribution. Moreover, the Central Limit Theorem rate of convergence $O(N^{-1/2})$ still applies. This justifies the idea of directly approximating the policy rather than the value functions. Clément et al. also show that taking $N^B \to \infty$ we recover the original value function $J^k(t, x, i)$.

4.3.4 Final Word On Convergence

Summarizing the last three sections, the total error we have is

$$\begin{aligned}\|Y_t^{k,i} - \hat{Y}^{k,i,\Delta,N}\|_2 &\leqslant \|Y_t^{k,i} - Y^{k,i,\Delta}\|_2 + \|Y_t^{k,i,\Delta} - \hat{Y}^{k,i,\Delta}\|_2 + \|\hat{Y}_t^{k,i,\Delta} - \hat{Y}^{k,i,\Delta,N}\|_2 \\ &\overset{?}{\leqslant} C\Big(\sqrt{\Delta t} + \frac{\mathbb{E}[\mathcal{R}_{t_1}(\mathbb{E}_{t_1}[Y_{t_2}^{1,i}])^2]}{\Delta t^k} + \frac{1}{\sqrt{\Delta t \cdot N}}\Big). \qquad (4.22)\end{aligned}$$

Even without the last term which is only our conjecture and is not proven, these convergence rates are quite pessimistic. We observe a general explosion in the regression errors as we cascade through the number of switches k, cf. the second term in (4.22). This would suggest that to obtain any convergence, the projection plus sampling errors must be decreasing exponentially in k. In the best case (with adaptive choice of (B_j)), we expect the projection error to decrease exponentially in the number of basis functions N^B, so the latter needs to grow linearly with k. However, an implicit issue that is hidden above is the relationship between N^B and N. Namely, for a given number of basis functions, how many Monte Carlo paths are needed to obtain a comparable sampling error. Unfortunately, Glasserman and Yu [41] give a negative result and show that in the worst case one may need as many as e^{N^B} paths. It follows that to compute $J^{\bar{K}}$ we must use at least $N \sim (\Delta t)^{-\bar{K}}$ paths, which would make the computations intractable.

In practice the LS scheme avoids repeated conditioning and therefore regression errors are accumulated only when a switch or a wrong switching decision is made. On a given path, instead of M regression errors, we should only see $O(1)$. As a result, we conjecture that the total error grows linearly in the maximum number of switches $\bar{K}$, which is what we seem to observe in practice. To prove this analytically,

one probably should use the fact that the barrier $\hat{\mathcal{M}}^{k,i}$ is closely related to the current $\hat{Y}^{k,i}$, and so the various errors in the barrier cancel out the errors in the regression approximation. This would produce tight bounds on expression of the form $\mathbb{E}[\hat{Y}^{k,i}_{t_2} \vee \hat{\mathcal{M}}^{k,i}_{t_1}]$ in terms of $\mathbb{E}[\hat{Y}^{k,i}_{t_2}]$ and $\mathbb{E}[\hat{\mathcal{M}}^{k,i}_{t_1}]$, cf. (4.19) and proof of (4.17).

From our extensive experimentation with the method, it seems that for $M < 1000$ the algorithm is quite stable. The corresponding discretization step Δt should be good enough for most financial problems. The real key to successful implementation is a correct estimate of the number of layers $\bar{K}$ and a judicious choice of the basis functions (B_j). After that $N \sim 20000$ simulated paths seem to be more than enough for examples we tried.

To check the actual rate of convergence with respect to the size of the Monte Carlo simulation we run Example 4.2 using $2000 - 24000$ paths and tabulate the resulting means and standard deviations over 50 separate runs in Table 4.1. We used six basis functions in this case and 400 time steps. We see that the variance decreases by about 40% when we double the number of simulations and then stabilizes as the remaining error is mostly coming from the projections and not from the Monte Carlo. We also see that in this example the value is decreasing, approaching the 'true' value of 5.931.

No. Paths	Mean	Std. Dev
2000	6.90	0.37
4000	6.44	0.30
8000	6.16	0.22
16000	5.90	0.19
24000	5.86	0.17

Table 4.1: Convergence of Monte Carlo simulations for Example 4.2.

4.4 Numerical Examples

We now give a series of numerical examples to illustrate the algorithm and the problem structure in different settings. Example 4.1 is the most basic one and has a one-dimensional time-homogeneous diffusion (X_t) with two regimes. Example 4.2 is the 'generic' one, with a two-dimensional diffusion (X_t) and three regimes. Example 4.3 shows the case where (X_t) has both jumps and seasonality. Finally, Example 4.4 shows a multi-dimensional, multi-mode setup with X_t taking values in $\mathbb{R}^3$ and $M = 5$.

Example 4.1. In our first example, we consider a simple one-dimensional Ornstein-Uhlenbeck driving process:

$$dX_t = 2(10 - X_t)\,dt + 2\,dW_t, \qquad X_0 = 10,$$

with time horizon $T = 2$, and delay (or rather switch separation) $\delta = 0.02$. We have two regimes with continuous reward rates of $\psi_0(X_t) = 0$ and $\psi_1(X_t) = 10(X_t - 10)$, and the switching cost between them is $C = 0.3$.

Figure 4.2 shows the output from the LS scheme for Example 4.1, using 200 steps, 8000 paths, and successive levels $k = 1, 2, \ldots, 8$. The initial regime is 'off', and the increasing curves correspond to the number of 'on-off' switches available. The plot shows each $J^k(t, x, 0; T)$ as a function of t, which due to the time homogeneity equals the expected total profit with horizon $T - t$. As we get further away from maturity, each separate value function J^k flattens out because the time decay becomes less relevant. The total value function $J = \sup_k J^k$ approaches linear growth in time to maturity since the stationary nature of (X_t) translates into an asymptotic fixed earnings per unit time ratio.

Figure 4.3 shows the switching thresholds in Example 4.1 as a function of time for two different levels k. The top boundary is for switching from regime 1 to regime 2 and the bottom is from regime 2 to regime 1. The area in between is the hysteresis

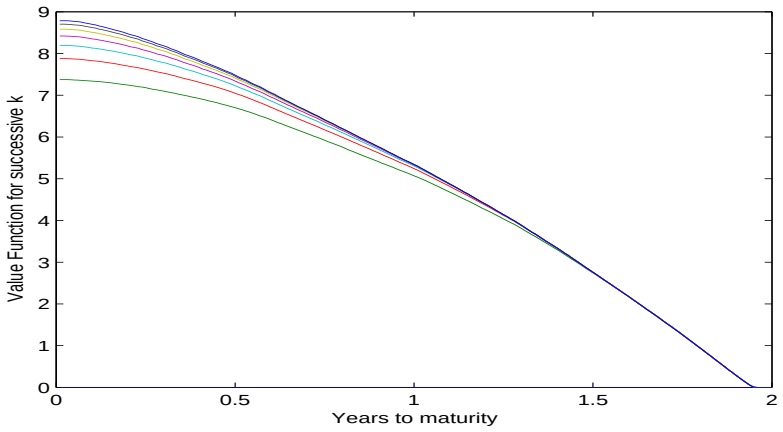
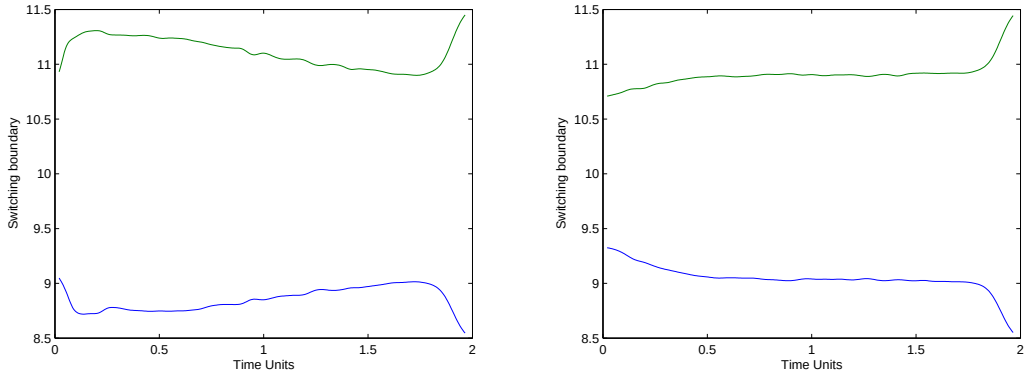


Figure 4.2: Value functions $J^k(t, x, 0)$ as a function of t for Example 4.1.

band where no switching takes place (note that at-the-money is always $x = 10$). Thus, if we the plant is 'off', it will be brought online only when X_t is about 10.8. The left panel plots the boundaries for $k = 2$. Far from maturity T the boundary is slowly increasing, as more time implies availability of more profitable switching opportunities. Closer to maturity, the time decay becomes significant and makes the agent less 'picky' in choosing the threshold to switch. Even a small expected profit is worthwhile to capture since little time is left. Finally, very close to maturity, the time decay dominates and the fixed cost is larger than any gain from switching. As a result, the switching boundary widens dramatically near T. For large k (the right panel in Figure 4.3), the exercise boundary is essentially flat after a sharp curve close to expiration, as time stationarity and large number of allowed switches make time decay insignificant. This parallels the situation with infinite horizon where the switching boundary is constant.

Example 4.2. The generic case of a spark spread, $X_t = (P_t, G_t)$ representing prices of power and gas respectively (for parameter meaning see (3.4)):

$$\begin{cases} \log(P_t) \sim OU(\kappa = 2, \theta = \log(10), \sigma = 0.8), & P_0 = 10, \\ \log(G_t) \sim OU(\kappa = 1, \theta = \log(10), \sigma = 0.4), & G_0 = 10, \qquad \rho_{pg} = 0.7. \end{cases}$$



$k = 1$ $k = 7$

Figure 4.3: Optimal Exercise Boundaries for Example 4.1 for different levels k. The decreasing boundary around $t = 0$ is an artifact of the Monte Carlo.

We solve the switching problem on $[0, 0.5]$ with regime rates of $\psi_0(X_t) = 0, \psi_1(X_t) = 10(P_t - G_t)$ and $\psi_2(X_t) = 20(P_t - 1.1\,G_t)$, and switching costs $C_{i,j} = 0.25|i - j|$. This example will be used for benchmarking in Section 4.6, see Table 4.4.

Example 4.3. Our third example shows that the method works also with jumps and time-dependent parameters. Let us consider the original spark-spread model with

$$d\log P_t = 6 \cdot (\bar{p}_t - \log P_t)\,dt + 0.5\,dW_t + \xi_t dN_t, \qquad P_0 = 10,$$
$$d\log G_t = 1 \cdot (10 - \log G_t)\,dt + 0.4\,dW_t^{\perp}, \qquad G_0 = 10, \qquad \rho_{pg} = 0.7,$$

with $\bar{p}_t = 10 \cdot [0.95 + 0.05\cos(2\pi t)]$ meant to represent seasonal fluctuation in electricity prices where the summer is consistently more expensive. (N_t) is an independent Poisson process with intensity $\lambda = 0.02$ and $\xi_t \sim exp(0.1)$, so that jumps are always upwards with a mean of 10% increase in price. There are two regimes again with the reward rates of $\psi_0(X_t) \equiv 0,\ \psi_1(X_t) = 10(P_t - G_t)$ and switching cost $C \equiv 0.5$.

For the seasonal Example 4.3, on the horizon of one year counting from July 1 to June 30th, it follows that the plant should rarely be operated in the winter around $t = 0.5$, when the mean power price is under \$9. However, as Figure 4.4 shows there are still sizeable gains to be made in the winter (over 10% of expected gains are generated in the three middle months). The possibility of price spikes or just large

enough fluctuations in (P_t) has value that can be exploited through plant flexibility. Of course, if large fixed costs can be avoided by complete shutdown it may still be optimal to close down the operations for the winter.

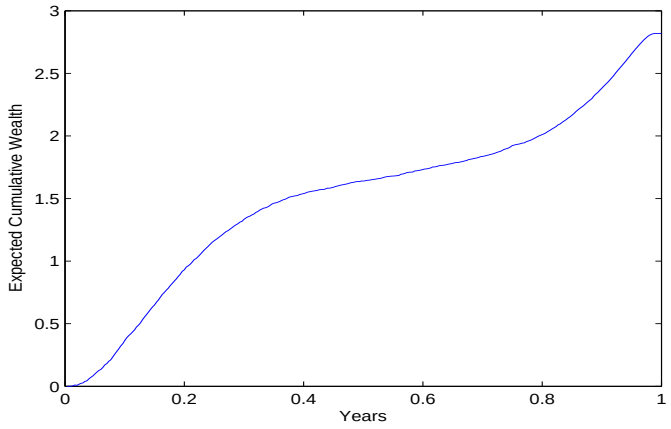


Figure 4.4: Value function for Example 4.3 showing expected cumulative profit $\mathbb{E}H(x, \text{'off'}, [0,t]; u^*)$ as a function of time t.

Example 4.4. The last example of this section illustrates the structure of optionality for a multi-commodity spread option. We consider a dual gas/crude oil power plant that can use either fuel depending on market conditions. The driving process is $X_t = (P_t, G_t, O_t)$ where (O_t) is the price of crude oil. Its dynamics are

$$\begin{cases} \log(P_t) \sim OU(\kappa = 2, \theta = \log(10), \sigma = 0.8), & P_0 = 10, \\ \log(G_t) \sim OU(\kappa = 1, \theta = \log(10), \sigma = 0.4), & G_0 = 10, \quad \rho_{pg} = 0.5, \\ \log(O_t) \sim OU(\kappa = 1, \theta = \log(10), \sigma = 0.4), & O_0 = 10, \quad \rho_{po} = 0.3, \ \rho_{go} = 0. \end{cases}$$

There are four possible regimes besides complete shutdown $\psi_0(X_t) \equiv 0$:

$$\psi_1(X_t) = 5 \cdot (P_t - G_t), \qquad \psi_2(X_t) = 5 \cdot (P_t - O_t),$$
$$\psi_3(X_t) = 5 \cdot (3P_t - 4G_t), \qquad \psi_4(X_t) = 5 \cdot (3P_t - 4O_t).$$

We take $C_{i,j} \equiv 0.25$ and also experiment with the delay setting, since switching a plant from one fuel to another is likely to take a long time. Hence we try the time-separation parameter $\delta = 0.01, 0.02, 0.04$ (up to two weeks) with $T = 1$ a horizon of

one year. As Table 4.2 shows, high δ significantly lowers the profitability, as does removal of regimes. For example, without regimes 3 and 4 which are initially 30% out of the money, the expected profit drops from 13.28 to 8.99. Similarly, if the plant could only run on gas, the expected profit nearly halves to just 7.06.

Regimes Avail.	Delay Setting	Value Function
All	PDE FD	13.31
All	$\delta = 0.01$	13.28
All	$\delta = 0.02$	12.49
All	$\delta = 0.04$	11.54
$0 - 3$	$\delta = 0.01$	10.41
$0 - 2$	$\delta = 0.01$	8.99
$0, 1, 3$	$\delta = 0.01$	7.06

Table 4.2: Summary of results from Example 4.4. We ran the LS scheme with 400 steps and 16000 paths. The first method is the finite differences PDE solver with no delay used for comparison.

Constructing Optimal Policies

Using the switching boundaries obtained from our algorithm, it is now straightforward to construct the optimal policy that can be implemented given a realization of (X_t). Suppose we computed $J^{\bar{K}}$ and our initial regime is u_0. Then we keep track of the switching sets $Switch_{u_0,j}$ for all $j \neq u_0$ corresponding to the $\bar{K}$-th level. At the first time t that $X_t \in Switch_{u_0,j}$ we switch to mode j. We then begin to monitor the switching boundaries of the $(\bar{K} - 1)$-th level and continue accordingly until T. It remains to keep track of the profits and losses along the way corresponding to the employed strategy. An illustration of this method for Example 4.1 is shown in Figure 4.5. The top panel shows the successive switching boundaries and a realized path of X_t. Observe the slight jumps of the boundaries at the switching times when the level changes. The bottom panel shows the total cumulative wealth obtained from the policy. The discrete switching costs are indicated by drops of $C = 0.3$ at the switching times.

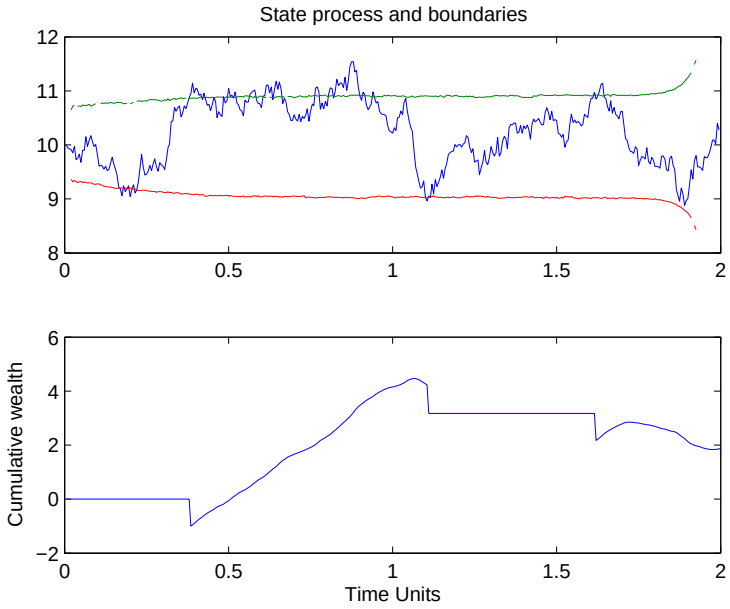


Figure 4.5: Implementing optimal policy for Example 4.1.

4.5 Comparison to Other Numerical Methods

The optimal switching problem allows for several other numerical methods of solution. Roughly speaking, there are three possible strategies. First, one can directly tackle the quasi-variational formulation of Section 3.6 using partial differential equation (PDE) solvers. Second, one can attempt to replace the continuous-space dynamics of (X_t) by some sort of discrete approximation. For this approximation one can then apply the dynamic programming approach directly. Finally, one can look for other means of computing conditional expectations in (4.1) besides the proposed Monte Carlo regression scheme. In this section we outline in turn each of these three approaches and finish by running a set of benchmarks to evaluate the relative performance and accuracy of our scheme.

PDE Solver

The PDE solver relies on the quasi-variational formulation of Proposition 3. The stochastic control problem is transformed into a parabolic partial differential equa-

tion with a free boundary. These have been heavily studied, and a variety of tools are available, see for example the references in Chapter 7 of Wilmott et al. [70]. For instance, consider the basic finite differencing (FD) algorithm. We begin by setting up a space-time grid (for simplicity take it to be uniform with steps Δt and Δx in time and space respectively). On this grid we solve

$$\begin{cases} u_t(t,x,i) + \mu(x)u'(t,x,i) + \frac{\sigma(x)^2}{2}u''(t,x,i) + \psi_i(x) = 0, \\ u(t,x,i) \geqslant \max_{j\neq i}(-C_{i,j} + u(t,x,j)), \\ u(T,x,i) = 0, \end{cases}$$

by replacing derivatives with finite differences in the first equation and directly enforcing the barrier condition at each step. Using the standard properties of the infinitesimal generator $\mathcal{L}_X$, $u(0,x,i) \xrightarrow{\Delta t\to 0, \Delta x\to 0} J(0,x,i)$. The free boundary can be determined by keeping track of the index j in the max argument of the barrier condition. An FD method is easy to implement but suffers several major drawbacks. First, numerical stability often requires taking a large number of time steps. Second, the method suffers from dimensionality problem: the size of the space grid is exponential in number of dimensions d and generally speaking $d > 3$ is computationally infeasible. Finally, the switching boundary will inevitably be jagged due to the presence of a grid in the x-space. The last point can be alleviated with the use of an adaptive grid. The method's accuracy depends on the order of the approximation used for the derivatives of u, which should be at least $O(\Delta t + \Delta x^2)$.

Markov Chain Approximation

The Markov chain approximation method pioneered by Kushner [52] consists in replacing (X_t) by a continuous time Markov chain $(\tilde{X}_t)$ with a finite state space $\tilde{E}$, such that its transition probabilities $(\tilde{P}_t)$ are consistent with the dynamics of (X_t). The final implementation is similar to the PDE solver since the transition probabilities are computed through the infinitesimal generator $\mathcal{L}_X$ of (X_t) which leads back to the finite differences approximations of the HJB equation.

For notational simplicity we will work with a one-dimensional (X_t). Take a rectangular grid $(x_n, t_m) = (x_0 + n\Delta x, m\Delta t)$, $0 \leqslant n \leqslant M^x$, $0 \leqslant m \leqslant M^\sharp$. Thus our finite base space is $\tilde{E} = \{x_0 + n\Delta x\} \subset E$ and appropriate boundary conditions will also be required. Next we assume that in one time step Δt, $\tilde{X}$ moves at most one point in the grid, meaning that the transition probability of going from $(x, m\Delta t)$ to $(y, m\Delta t + \Delta t)$, which we denote by $P_m(x, y)$ is zero unless $y \in \{x, x + \Delta x, x - \Delta x\}$. To solve for P_m we need consistency conditions which make the dynamics of $(\tilde{X}_t)$ match the dynamics of (X_t). If we match the first two moments, then

$$\begin{cases} \mathbb{E}[\tilde{X}_{(m+1)\Delta t} | \, \tilde{X}_{m\Delta t} = x] - x = \mu(x)\Delta t \quad \text{and} \\ Var(\tilde{X}_{(m+1)\Delta t} | \, \tilde{X}_{m\Delta t} = x) = \sigma^2(x)\Delta t. \end{cases}$$

Next we must decide whether P_m is given explicitly or implicitly. For example, the upwind implicit version for computing the conditional expectation $\hat{J}^k(t_1, x, i) = \mathbb{E}[J^k(t_2, X_{t_2}, i) | X_{t_1} = x]$ is:

$$\Big[1 + \sigma^2(x)\frac{\Delta t}{\Delta x^2} + |\mu(x)|\frac{\Delta t}{\Delta x}\Big]\hat{J}^k(t_1, x, i) = \Big[\frac{\sigma^2(x)}{2}\frac{\Delta t}{\Delta x^2} + \mu^+(x)\frac{\Delta t}{\Delta x}\Big]\hat{J}^k(t_1, x + \Delta x, i) + \Big[\frac{\sigma^2(x)}{2}\frac{\Delta t}{\Delta x^2} + \mu^-(x)\frac{\Delta t}{\Delta x}\Big]\hat{J}^k(t_1, x - \Delta x, i) + J^k(t_2, x, i), \quad (4.23)$$

where $x^+ \triangleq \max(x, 0)$ and $x^- \triangleq -\max(-x, 0))$. This can be formally re-written as

$$\hat{J}^k(m\Delta t, x, i) = \sum_{y \in \tilde{E}} P_m(x, y) \cdot \hat{J}^k(m\Delta t, y, i) + P_m(x, x) \cdot J^k((m+1)\Delta t, x, i)$$

indicating the dynamic programming analogy. Solving such an implicit system only involves inverting a tri-diagonal matrix. Thus the code is $O(1/\Delta t)$ fast. After $\hat{J}(m\Delta t, x_0 + n\Delta x, i)$ is solved for, we simply mimic (4.1):

$$J^k(m\Delta t, x_0 + n\Delta x, i) = \Big(\hat{J}^k(m\Delta t, x_0 + n\Delta x, i) + \psi_i(m\Delta t, x_0 + n\Delta x)\Delta t\Big) \vee \max_{j \neq i}\Big(J^{k-1}(m\Delta t, x_0 + n\Delta x, j) - C_{i,j}\Big).$$

Like the PDE solver, this algorithm suffers from the curse of dimensionality and is not robust to extensions. Nevertheless, in low dimensions it provides a reliable

benchmark. We have implemented the basic explicit scheme for dimensions one through three, and the implicit scheme in one dimension.

Quantization Method

A powerful non-Markovian version of the above method with an adaptive approximating grid is the so-called Quantization Scheme. The main motivation of quantization is to find a small and efficient approximating grid in exchange for giving up the Markov property and closed-form formulae for P_m. The latter is instead computed via a Monte Carlo simulation. While this is likely to be slow, it can be done just once off-line and stored for later calculations. The gain is increased robustness and much better dimensional scaling.

To implement the quantization method we need two major ingredients: the state spaces $\tilde{E} = \{\tilde{E}_m\}_{m=0}^{M\sharp}$ and the transition matrices $\tilde{P} = \{P_m\}$. The location sites $\tilde{E}_m = \{x_m^k\}, k = 1, \ldots, N_m$, called the quantization grids, should be a cloud of N_m points optimal in the sense of approximating well the distribution of $X_{m\Delta t}$. More precisely, we want $\mathbb{E}\big[\min_k \|X_{m\Delta t} - x_m^k\|^2\big]$ to be small. The proper way to think of the quantization grid is as a partition of $E \subseteq \mathbb{R}^d$ into cells. The cells are the Voronoi tessellations $\{C_m^k\}$ such that $C_m^k \triangleq \{x \in \mathbb{R}^d \colon \|x - x_m^k\| = \min_k \|x - x_m^k\|\}$ is the set of all x's closest to x_m^k. The transition matrices P_m are then required to approximate closely the dynamics of (X_t) between cells of the adjacent quantization grids. Namely, we would like

$$P_m(i,j) \simeq \mathbb{P}\left[X_{(m+1)\Delta t} \in C_{m+1}^j \,\middle|\, X_{m\Delta t} \in C_m^i\right].$$

Note that the quantized process $\tilde{X}_{m\Delta t} = \sum_k x_m^k \cdot \mathbb{1}_{C_m^k}(X_{m\Delta t})$ obtained by picking the nearest neighbor on the quantization grid, is *not* Markov, so the above expression is ill-defined and we really should be conditioning on the entire path of (X_t). Algorithmically $(\tilde{E}, \tilde{P})$ is constructed by simulation. In particular, Bally et al. [5] propose the Competitive Learning Vector Quantizer (CLVQ) algorithm for dynamically con-

structing the grid. The basic concept is to simulate N paths $(x^{\ell}_{m\Delta t}, \ell = 1, \dots, N)$ of (X_t) and at each stage to pick the 'winner' indices $i_m(\ell)$ satisfying $x^{\ell}_{m\Delta t} \in C^{i_m}_m$. Now we simply approximate with the empirical average

$$P_m(i,j) = \frac{\text{card}\left|x^{\ell}_{m\Delta t} \in C^i_m, x^{\ell}_{(m+1)\Delta t} \in C^j_{m+1}\right|}{\text{card}\left|x^{\ell}_{m\Delta t} \in C^i_m\right|}. \tag{4.24}$$

The CLVQ further improves by slightly adjusting the grid $\tilde{E}_m$ given the simulated paths in order to obtain the best possible (in-sample) quantizing grid. Bally et al. [5] show that the resulting speed of convergence to the true Snell envelope in a standard optimal stopping problem is $O\Big(\frac{(\Delta t)^{1+d}}{\sum_m N_m}\Big)$, and the L^p error of the quantization grid itself is $O((\sum_m N_m)^{-1/d})$. They also show that there is a gain from using variable grid sizes N_m with smaller grids towards $t = 0$ and larger grids around $t = T$. Heuristically, a grid size of $500 - 1000$ points is usually sufficient in dimension $d < 3$. Figure 4.6 shows an optimal quantization grid for the 2-d exponential OU process of Example 4.2 using 600 sites.

Once $(\tilde{E}, \tilde{P})$ is computed, we solve the optimal switching problem by formally computing the pseudo-Snell envelopes of the non-Markov $(\tilde{X}_t)$ using the familiar recursion (both sides are vectors now):

$$J^k(t_1, \tilde{E}_{t_1}, i) = \max\Big(\psi_i(t_1, \tilde{E}_{t_1})\,\Delta t + J^k(t_2, \tilde{E}_{t_2}, i) \cdot P'_{t_1}, \max_{j \neq i}\big\{-C|i-j| + J^{k-1}(t_1, \tilde{E}_{t_1}, j)\big\}\Big). \tag{4.25}$$

The advantage of the quantization method is its robustness, since the entire algorithm for constructing $(\tilde{X}_t)$ is Monte Carlo based. Moreover, the simulations themselves are straightforward and can be parallelized. Once the grids and transitions are computed, *any* control problem can be solved because the approximation of (X_t) is independent from the optimization step. Contrast this with the regression method where the basis functions are customized to the payoff functions ψ_i.

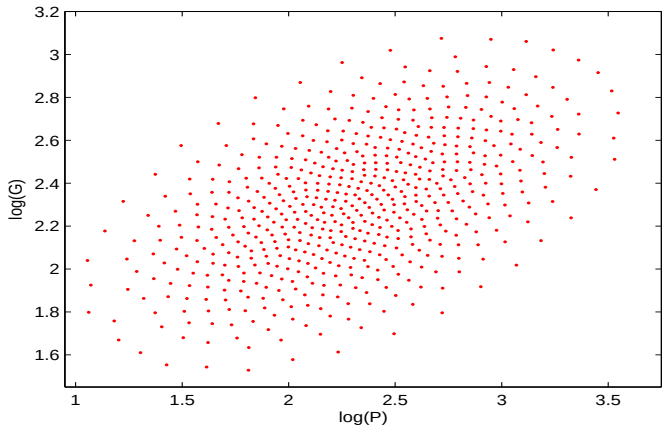


Figure 4.6: Example of optimal quantization grid for Example 4.2. We use 600 points and the grid for standard 2-d Brownian motion from Gilles Pagés webpage [63].

The difficulty with quantization is estimating $\tilde{P}$. When there are many cells, each C_m^k shrinks, causing the estimate in (4.24) to have high variance since both the numerator and denominator are very small. Thus, one is forced to run a huge number of simulations (often as many as 10^6) to achieve an acceptable accuracy. One solution, proposed by Mrad et al. [59][3] is to treat C_m^k as a single point and to approximate the joint density in (4.24) by using Malliavin integration by parts. The authors claim a great numerical improvement, but further analysis is needed to understand when it is acceptable to replace the cell C_m^k by a Dirac mass at x_m^k.

Nonparametric regressions

The final alternative we discuss concerns directly our own method of discretizing the Snell envelope. As we mentioned before, this approach reduces to computations of conditional expectations. The latter can be obtained by several methods, including Malliavin calculus [11], Monte Carlo simulations and regression. In particular, among regression schemes we still have a choice between regression against basis functions that we performed versus fully non-parametric regression. Choosing a

[3]I am grateful to Nizar Touzi for pointing out this reference to me.

non-parametric regression relieves us of the concerns regarding selecting appropriate basis functions and may produce smoother conditional distributions [64].

We have explored the simplest version of this approach, namely k-nearest neighbors multivariate kernel regression. Given a set of simulated paths $(x^\ell_{m\Delta t})$ like in Section 4.1 and setting $y^\ell_{m+1} \triangleq J((m+1)\Delta t, x^\ell_{(m+1)\Delta t}, i)$ we approximate

$$x \mapsto \mathbb{E}\Big[J((m+1)\Delta t, X_{(m+1)\Delta t}, i)\Big|\, X_{m\Delta t} = x\Big] \quad \text{by} \quad x^\ell_{m\Delta t} \mapsto \sum_{i=1}^{k} w_{\alpha_i} y^\ell_{m+1},$$

where α_i are the k nearest neighbors of $x^\ell_{m\Delta t} \in \mathbb{R}^d$ and

$$w_{\alpha_i} = \frac{K(\frac{1}{h}\cdot(x^\ell_{m\Delta t} - x^{\alpha_i}_{m\Delta t}))}{\sum_{i=1}^k K(\frac{1}{h}\cdot(x^\ell_{m\Delta t} - x^{\alpha_i}_{m\Delta t}))}.$$

The kernel $K\colon \mathbb{R}^d \to \mathbb{R}_+$ is a smooth $\mathcal{C}^\infty$ function which integrates to one and is centered at zero. For instance, a classical choice is the Gaussian kernel,

$$K(x) = \frac{1}{\sqrt{2\pi}^d}\, \mathrm{e}^{-\|x\|^2/2}.$$

Thus, the regression is replaced by a local linear combination of the other paths' values with the weights proportional to the distance. The use of nearest neighbors is to reduce the curse of dimensionality. The simple kernel regression algorithm that uses all points would have complexity $O(N^2)$. By contrast, the complexity of the k-nearest neighbors version is $O(N \log N + kN)$ since we need to compute N sums of k terms, and to select those k terms we need to sort once the whole data set. Hence taking $k \sim \log N$ we only do $O(N \log N)$ work. In more than one dimension selecting nearest k neighbors in $O(N \log N)$ time is challenging. A simple scheme is to pick a variable number of neighbors instead. We first sort over one coordinate, and then define an adaptive hypercube for the neighbors in the other coordinates. Namely we fix parameters N_1 and D and for a given site $x \equiv (x_1, \ldots, x_d) \in \mathbb{R}^d$ pick

$$Nbr(x) = \{x^\ell\colon |x^\ell_1 - x_1| \text{ is among } N_1 \text{ smallest values}, |x^\ell_j - x_j| \leqslant D \text{ for } j = 2, \ldots, d\}.$$

The kernel estimator converges to the true conditional expectation with the speed of $O(N^{-1/2})$, however it is usually biased [19]. It is attractive due to its robustness, since besides averaging no additional error is introduced. However, the two major difficulties with the kernel method are selecting an appropriate bandwidth and the computations around the edges. The bandwidth h controls the peakedness of the weights around (x, y). Thus, as the bandwidth increases, more distant points carry more weight and the estimate becomes more 'global'. However, choosing h is heuristic and may require a lot of trial-and-error. The other difficulty is when the regressed y^ℓ is extreme. Consider, for example, the situation when $y^\ell_{m+1} = \max_i y^i_{m+1}$. Then all the neighbors of $(x^\ell_{m\Delta t}, y^\ell_{m+1})$ have smaller y-values causing the estimate at $x^\ell_{m\Delta t}$ to be necessarily underestimated. Observe that by contrast, the regression algorithm does not have any inherent bias for extreme response values.

4.6 Benchmarking

As mentioned in introduction, we do not know of any papers that have solved a switching problem on a finite horizon. Therefore, to verify our numerical results we initially benchmarked our scheme in the easier American option setting. More precisely, we adjusted our algorithms to solve the 2-dimensional American exchange option that is discussed in Villeneuve and Zanette [68].

The reference problem is pricing the American put option paying out $(K - \min(S^1_\tau, S^2_\tau))^+$ where (S^1, S^2) follow the two-dimensional Black-Scholes model

$$\begin{cases} dS^1_t &= \log(1.05)\, dt + 0.3\, dW^1_t, \\ dS^2_t &= \log(1.05)\, dt + 0.2\, dW^2_t \end{cases}$$

with $d\langle W^1, W^2\rangle_t = 0.5\, dt$. The option is at the money with $K = S^1_0 = S^2_0 = 40$ and the maturity is $T = 7/12$, seven months. The reference value is 3.8958 from [68, p. 147]. We recomputed this option value using our homegrown versions of the Longstaff Schwartz (LS), Tsitsiklis van Roy (TvR), Kernel, Quantization and

Markov Chain approximation (i.e. PDE FD) methods. The results are in Table 4.3. All the regression methods used 400 time steps and 10000 paths. The PDE solver used 20000 steps and a 400×400 grid in space. For the quantization method we used $N = 250{,}000$ and grids with sizes up to $N_m = 600$ taken from [63]. The variance was obtained by running each algorithm 50 times and the timing is relative to a Pentium 4, 1.8GHz desktop.

On this simple problem all methods did very well. We see that the PDE solver did the best, exactly reproducing the reference value. However, it was quite slow. The second best was surprisingly the LS scheme. The kernel algorithm was quite erratic and required a lot of finetuning. The quantization did not perform very well, however we have not finetuned it at all and a smarter implementation should do much better. Note that the given time of nearly seven hours is for computing all the transition probabilities (P_m). However, once they are computed any optimal stopping problem for (S^1, S^2) can be solved almost instantaneously.

Method	Mean	Std. Dev	Time (m)
Explicit FD	3.8955	–	7.5
LS Regression	3.929	0.03	0.8
TvR Regression	4.213	0.02	0.8
Kernel	4.078	0.06	3.1
Quantization	4.025	0.01	400*

Table 4.3: Benchmark results for the two-dimensional American minimum put option. The quantization method time is for computing $\tilde{P}$. See comment in text.

We proceed to benchmark our computation of Example 4.2. We choose this case because it is two-dimensional and has three regimes, which can be viewed as a generic setting of our problem. Again, we repeat the results using all methods listed above. In absence of third party reference value we assume that the PDE solver is the most accurate since its error for this problem should be comparable to the error on the American put which was negligible. This time the kernel method performed best, however we believe this is just a lucky coincidence. The LS scheme did very well,

though its variance was quite high. In contrast, the TvR scheme had smaller variance but a very strong bias, confirming the superiority of the Longstaff-Schwartz approach. The quantization performed acceptably and again a better implementation should make it competitive.

Summarizing, in small dimensions ($d < 3$) the best algorithm is the PDE solver. An industrial-strength implementation should be very fast and produce provably accurate results. However, the PDE approach becomes infeasible as soon as $d > 2$, as grid sizes approach 10^6 points. Thus, for scalability to high dimensional problems the LS and Quantization schemes should be used. The choice between those depends on the environment. If one is looking for a quick tool to solve optimal switching for a variety of (X_t), the LS scheme is best. On the other hand, if (X_t) is fixed, the quantization may be better as one has time to compute a superior approximation of $\tilde{P}$ and then use it to price any (X_t)-tolling agreements within seconds.

Method	Mean	Std. Dev	Time (m)
Explicit FD	5.931	–	25
LS Regression	5.903	0.165	1.46
TvR Regression	5.276	0.096	1.45
Kernel	5.916	0.074	3.8
Quantization	5.658	0.013	400*

Table 4.4: Benchmark results for Example 4.2. Computational settings are same as for Table 4.3.

4.7 Comparison to Practitioner Methods

Let us briefly summarize the tools that have been used by practitioners to solve the operational flexibility problem. Early on the standard approach was the classical net present value theory. It used discounted cash flow analysis to estimate the value of the asset based on projections of future prices and proper weighing and discounting of possible cases. Uncertainty was essentially eliminated, as static scenarios were used to forecast the future and select pre-determined optimal behavior. The opportunity

of dynamically responding to prices was ignored and as a result the contracts were consistently underpriced [26].

In late 1980s, the stochastic programming approach, also called Markov Decision Processes (MDP) became popular. The terminology of MDP's is rather different, but in effect they are tree-based versions of the stochastic control formulation. The problem is discretized in time and the path space of (X_t) is broken into a finite number of scenarios. Then the transition probabilities $P(x, y; \Delta t)$ are estimated (by simulation or analytical approximation) for each current outcome x and possible transition scenario y. Finally, the problem is solved via dynamic programming that corresponds to a lattice discretization of the QVI :

$$J(t_1, x, i) = \max_j \big\{ -C_{i,j} + \sum_y P_{t_1}(x, y) \cdot J(t_2, y, j) \big\}.$$

Like with American options, the MDP tree is simple to write down and intuitive to understand. An example of explicit solution for a simple switching problem can be found in [71]. However, if one must solve numerically then the computational complexity explodes for long horizons with many optionalities.

With the advent of financial engineering, the new widely used method for pricing tolling agreements is the strips of spark-spread options approach [38]. The motivation is to reduce the problem to pricing standard financial options whose valuation is well understood. Accordingly, the payoff from the power plant is represented as a collection of European options that pay the maximum value to be obtained during each period. Each such option is of the spark-spread variety with maturity $T_m = m\Delta t$. If for simplicity we consider just two regimes 'off' and 'on', the overall value of running the power plant is approximated by

$$V(0, x; \Delta t) = \mathbb{E}^x \Big[\sum_{m=0}^{M^\sharp} \pi_{m\Delta t} (P_{m\Delta t} - \overline{HR} \cdot G_{m\Delta t} - K) \cdot \Delta t \Big], \quad (P_0, G_0) = x. \quad (4.26)$$

Above $\pi_{m\Delta t}$ is the exercise decision variable, equivalent to the policy u used in the control framework. In the no-delay setting we allow the plant to be switched

instantaneously, so that $\pi_t = \mathbb{1}_{P_t > \overline{HR} \cdot G_t + K}$ producing the classical spread option. With delay we require $\pi_t \in \mathcal{F}_{t-\delta}$ to be predictable beforehand, so the option expires before maturity.

Let us stress that the strip approximation introduces two major simplifications. First, it ignores the time-dependency of the optimal policy and the hysteresis band resulting from presence of switching costs. Second, it eliminates the time decay due to finite-horizon features of the contract[4] . Nevertheless, the method remains immensely popular. One reason is that for many processes (X_t) (in particular the conditionally Gaussian case) there are efficient closed form approximations to spread options that are very fast to compute and provide good bounds not just on the price but also on the sensitivities [15]. As a result, the general intuition that practitioners have built up for dealing with vanilla options can be directly transferred.

For direct comparison of the two valuations, the major obstacle is the switching cost $C_{i,j}$. Indeed, as the next proposition shows, without switching costs the two approaches are equivalent.

Lemma 4. *Let $C = \sup_{i,j} C_{i,j}$. In the limit of vanishing switching costs $C \to 0$ the discretized value function $J(0, x, i)$ of optimal switching converges to the value function $V(0, x)$ coming from the spark spread approximation.*

Proof. Fix a time step Δt and current time $t_1 = m\Delta t$. For simplicity we work without delay and with $K = 0$, in which case

$$V(m\Delta t, x) = \mathbb{E}\Big[\sum_{j=m}^{M^\sharp} (P_{j\Delta t} - \overline{HR} \cdot G_{j\Delta t})^+ \cdot \Delta t \Big| \, X_{m\Delta t} = x\Big].$$

[4]To account for some of the operational constraints other variants have also been considered, for instance swing options [17] that limit the total number of exercises.

Clearly, $V(t_1, x)$ is the maximum value that one can extract from the plant since there are no operational constraints, so trivially $V(t_1, x) \geqslant J(t_1, x, i)$. By a basic conditioning,

$$V(t_1, x) = \mathbb{E}[V(t_2, X_{t_2}^{t,x}) | \, X_{t_1} = x] + (P_{t_1} - \overline{HR} \cdot G_{t_1})^+ \cdot \Delta t. \tag{4.27}$$

On the other hand, for any x there exists mode i such that $\psi_i(x) = (P_{t_1} - \overline{HR} \cdot G_{t_1})^+$. For this mode i,

$$J(t_1, x, i) \geqslant \mathbb{E}\big[J(t_2, X_{t_2}^{t,x}, i) | \, X_{t_1} = x\big] + \psi_i(x) \Delta t,$$

and combining with (4.27),

$$|J(t_1, x, i) - V(t_1, x)| \leqslant \mathbb{E}\big[|J(t_2, X_{t_2}^{t,x}, i) - V(t_2, X_{t_2}^{t,x})| \, \big| \, X_{t_1} = x\big]$$

Now we use the fact that $|J(t, x, i) - J(t, x, j)| \leqslant C$ to get

$$\sup_x \max_i |J(t_1, x, i) - V(t_1, x)| \leqslant C + \sup_x \max_i |J(t_2, x, i) - V(t_2, x)|.$$

Inducting on t_1, we get $\sup_x \sup_t |J(t, x, i) - V(t, x)| \leqslant M^\sharp \cdot C$ and taking the limit $C \to 0$ we are done. □

Returning to the case when $C_{i,j}$'s are significant, we observe that with the strip of options approach $C_{i,j}$ does not enter into the decision process so that many exercises are made to capture relatively small gains. One solution to correct for this could be to increase the strike price of the option to include the switching cost. However, this still ignores the time-series properties of running the plant which form a significant component of the total price. Indeed, once we switch 'on', there is a high probability of being in-the-money in the next time period as well. A better fix is first to compute the unadjusted $V(t, x)$ and then estimate the corresponding number of switches. For instance, assume that we switch off immediately after reaching zero profitability, and switch on as soon as the spread is at least $P_t - \overline{HR} \cdot G_t > a$ for some threshold $a > 0$.

Ideally a is equal to the true switching boundary, but any rough estimate would do. Then the approximate number of switches is equal to the number of downcrossings by $P_t - \overline{HR} \cdot G_t$ over the interval $[a, 0]$. Each downcrossing corresponds to an 'on-off' control and therefore costs $C_{0,1} + C_{1,0}$ that we subtract from $V(t, x)$. In reality profitable paths will remain in-the-money for a long time and therefore will have few downcrossings, while the paths with many downcrossings will always stay close to at-the-money and therefore would not be worthwhile to exercise. In this light, the adjusted strip estimate is a lower bound, since it will overestimate the switching costs.

For a one-dimensional Orstein-Uhlenbeck process, the distribution of downcrossings may be analytically computed. Using the explicit formula for Laplace transform of a hitting time we have [39]

$$\mathbb{E}[T_{0,a}^{down}] = \sqrt{\frac{\pi}{2}} \int_{-a}^{a} (1 + \operatorname{erf}(t/\sqrt{2})) \cdot \mathrm{e}^{t^2/2}\, dt,$$
$$Var[T_{0,a}^{down}] = \sqrt{2\pi} \int_{-a}^{a} \int_{-\infty}^{t} \int_{s}^{a\mathbb{1}_{\{t>0\}}} (1 + \operatorname{erf}(r/\sqrt{2})) \cdot \mathrm{e}^{(r^2+t^2-s^2)/2}\, dr\, ds\, dt,$$

where $T_{0,a}^{down}$ is the time for a standard OU process to complete one downcrossing across $(a, 0)$. For the range of interest this shows that the expected number of downcrossings is $O(a^{-3/2})$. To illustrate, consider Example 4.1 with switching cost of $C = 1$. The optimal value for the stochastic control model comes out to be about $J(0, 10, \text{off}) = 4.23$. On the other hand, the expected profit from a strip of spread options is $V(0, 10) = 7.65$. Using as our guide the exercise boundary in Figure 4.3, a downcrossing interval of $[10, 11]$ yields 1.93 switches on average for a final profit of 3.8, which is within 10% of the true value. In more general situations, the number of downcrossings can be estimated by a simple Monte Carlo simulation.

Remark 3. Analysis of downcrossings of (X_t) can also shed light on the error between J^k and the true value function J. Indeed, the probability that more than k switches are needed for optimal strategy can be bounded from above by some conservative

estimate on the number of 'profitable switching opportunities' which are closely related to (X_t) crossing the switching boundaries. Of course the true switching boundary is unknown but we can get a good estimate from the boundaries of J^k.

4.8 Simulating Mean-Reverting Processes

All the numerical algorithms described in this chapter require simulation of (X_t) so let us address this issue briefly. We first concentrate on the 'classical' setup from (3.4), where the pair of price processes (G_t, P_t) follow a two-dimensional exponential Ornstein Uhlenbeck process. For convenience we will work with their logarithms $(\bar{P}_t, \bar{G}_t) \equiv (\log P_t, \log G_t)$ which therefore satisfy

$$\begin{aligned} d\bar{P}_t &= \kappa_p(\bar{\theta}_p - \bar{P}_t)\,dt + \sigma_p\,dW_t^1, \\ d\bar{G}_t &= \kappa_g(\bar{\theta}_g - \bar{G}_t)\,dt + \sigma_g\,dW_t^2, \end{aligned}$$

with $d\langle W^1, W^2\rangle_t = \rho\,dt$ and $\bar{\theta}_i = \theta_i - \frac{\sigma_i^2}{2\kappa_i}$. Recall that given an initial condition (g_0, p_0), the process is Gaussian and the conditional mean and covariance of $(\bar{P}_t, \bar{G}_t)$ are explicitly given by [15]

$$\begin{aligned} \mathbb{E}[\bar{P}_t|\,P_s, G_s] &= \bar{\theta}_p + \mathrm{e}^{-\kappa_p(t-s)}(\log P_s - \bar{\theta}_p), \\ Var[\bar{P}_t|\,P_s, G_s] &= \sigma_p^2\,\frac{(1-\mathrm{e}^{-2\kappa_p(t-s)})}{2\kappa_p} \qquad \text{similarly for } \bar{G}_t, \\ \mathbb{E}\big[\bar{P}_t\cdot\bar{G}_t|\,P_s, G_s\big] &= \rho\sigma_p\sigma_g\,\frac{1-\exp(-(\kappa_g+\kappa_p)(t-s))}{\kappa_g+\kappa_p}. \end{aligned}$$

This implies that (P_t, G_t) are conditionally log-normal and the process can be simulated *exactly*, eliminating SDE discretization error when computing $X_{m\Delta t}$. Note that the explicit formulae pertain even when the mean reversion level is time-dependent, e.g. $\theta_p = \theta_p(t)$. In this case,

$$\mathbb{E}\big[\bar{P}_t\big] = \int_0^t \kappa_p \mathrm{e}^{\kappa_p(s-t)}(\theta_p(s) - \frac{\sigma_p^2}{2\kappa_p})\,ds + \mathrm{e}^{-\kappa_p t}\log P_0,$$

with same variance as before.

Mean Reversion with Jumps

We can also easily incorporate Poisson jumps with deterministic intensity in either P_t or G_t. This means that the SDE for, say, P_t has an extra term dN_t, where N_t is a Poisson process with intensity λ, cf. Example 4.3. To simulate $P_{t+\Delta t}$ given P_t, first simulate the number of jumps on $[t, t+\Delta t]$. Then conditional on $N_{t+\Delta t} - N_t$, the times of the jumps are uniformly distributed with respect to the cumulative intensity process. It remains to simulate the size of each jump (which are assumed to be i.i.d.) and let the process follow the Gaussian OU dynamics between jump times.

On a practical level the problem is that to achieve the large spikes prevalent in observed power prices the mean-reversion parameter has to be very large. Usually the resulting estimates of κ_p are unacceptably large, killing off any volatility besides the jumps. One solution [50], is to use an additive two-factor model for $P_t = P_{1\,t} + P_{2\,t}$, with one factor $P_{1\,t}$ an OU diffusion as before, and a second mean-reverting jump factor $P_{2\,t}$ driven entirely by a Poisson process which is responsible for the jumps. The attractive feature of such a model is that one can still simulate exactly and moreover derive closed-form formulae for vanilla options. For further details about modeling spot prices of commodities we refer to [16] and the references therein.

Chapter 5

Generalizations

5.1 Gas Storage

As opposed to purely financial obligations like stocks and bonds, commodities require physical storage and consequently often exhibit seasonal price patterns. This is especially noticeable in energy fuel commodities whose consumption is highly correlated with weather. The principal example is natural gas that is used by many households in North America for heating during the winter. Thus, natural gas demand has a pronounced spike in the cold season. On the other hand, natural gas supply, obtained by extracting gas from gas fields and moving it through the pipeline system closer to the end users, is relatively stable. To accommodate the higher winter demand there are all sorts of storage facilities, like salt domes, depleted gas fields and aquifers that allow for gas to be kept in an easily accessible manner.

Given that short term gas prices are highly volatile and unpredictable, storage also provides a tool for financial speculation. In particular, salt domes have high deliverability rate that allows to make speculative bets on intertemporal spreads in gas prices. As documented by de Jong and Walet [25], "in the liberalisation process, natural gas storage is unbundled, ... offered as a distinct, separately charged service. ... Buyers and sellers of natural gas have the possibility to use storage capacity to take

advantages of the volatility in prices". The basic idea is to rent a storage facility and then to 'buy low' and 'sell high', such that the realized profit covers the intermediate storage and rental costs. This is in close parallel to power plant tolling discussed in the first half of this dissertation. Continuing the analogy, it is now natural to ask what is the financial value of such a rent contract. Observe that the time structure and seasonality of prices becomes crucial. As a stylized example, if average gas prices are \$4 per Btu in July and \$8 in January, while monthly storage costs are $50c$ cents, then one can realize a sure profit of \$1 by buying gas in July, forward selling it in January and storing in between. Such transactions do take place, but they should be seen as monetizing the economic rent of the storage facility rather than arbitrage. Indeed, the total amount of storage is limited and the owner has to invest very large amount of capital and time to build the facility in the first place. We are more interested instead in dynamic trading that responds to short-term fluctuations. For us a gas storage facility acts like a straddle on gas prices, making the agent long volatility. Thus, in this presentation we will ignore the forward curve dynamics and focus on the 'real option' component of storing gas which derives value from the timing optionality.

To construct a tractable mathematical model we assume that the storage facility operates as follows. First, gas is transferred in and out by means of a pipeline which has limited capacity, implying that the transfer rate is fixed by the operational characteristics. For simplicity, assume that we have a single feasible injection rate a_{in} and a single withdrawal rate a_{out} in units of MMBtu per time unit. Furthermore, assume we have a constant storage rate b for storing one MMBtu for one time period. Finally, assume that changing the operating regime of the facility is costly and costs C per switch. This cost represents both the effort—one must dispatch workers, coordinate with the outgoing pipeline, stop/start the decompressors, etc.—and the time needed to change the operating mode. In practice, storage is subject to many

other operational constraints. These include upper and lower bounds on inventory, injection and withdrawal constraints, seepage costs, etc. All these may be time- and inventory level-dependent, but for now we simply ignore them.

For gas prices we mainly refer to near-month forwards that are by far the most liquid contract on the market, however overall we remain agnostic about the interpretation of the price process. In any case, letting G_t denote the current gas price and I_t the current inventory in storage, the facility can then be in three possible regimes $i \in \{-1, 0, 1\}$ with corresponding payoff rates given by

$$\begin{cases} \text{Inject:} & \psi_{-1}(G_t, I_t) &= -G_t \cdot a_{in} - b\, I_t, & dI_t &= a_{in}\, dt, \\ \text{Store:} & \psi_0(G_t, I_t) &= -b\, I_t, & dI_t &= 0, \\ \text{Withdraw:} & \psi_1(G_t, I_t) &= +G_t \cdot a_{out} - b\, I_t, & dI_t &= -a_{out}\, dt. \end{cases} \tag{5.1}$$

The facility operator would then like to maximize the net profit given that changing regimes and storage are costly:

$$J(t, g, c, i) \triangleq \sup_{u \in \mathcal{U}(t)} \mathbb{E}\left[\int_t^T \psi_{u_s}(G_s, I_s)\, ds - C \int_t^T |du|_s \,\middle|\, G_t = g, I_t = c, u_t = i \right]. \tag{5.2}$$

The contract normally specifies that the facility should be returned with the same inventory I_0 as in the beginning. Various buy-back provisions are employed to enforce violations of this constraint at the end of the storage period. A possible terminal condition is $J(T, g, I_T, i; I_0) = -\bar{C} \cdot g|I_T - I_0|$, making the penalty proportional to the absolute difference with stipulated inventory times the terminal gas price. One will often also require a specific final mode of the facility, say $u_T = 0$.

The control problem (5.2) is somewhat different because now the dynamics of (X_t), namely the inventory I_t, is affected by the control. Nevertheless, we can reuse the same methodology as in Section 3.3. Define the recursive value functions

$J^k(t,g,c,i)$ to denote the optimal expected profit to be had starting at time t with $G_t = g, I_t = c$ while in state $u_t = i$ and having $k = 0, 1, \ldots$ switches left:

$$J^k(t,g,c,i) = \sup_{\tau\in\mathcal{S}_t} \mathbb{E}\Big[\int_t^\tau \psi_i(G_s, I_s)\, ds + \max_{j\neq i}\big\{-C_{i,j} + J^{k-1}(\tau, G_\tau, I_\tau, j)\big\}\Big| G_t = g, I_t = c\Big]. \quad (5.3)$$

Similar to (3.10), we see that the optimal switching times τ^* should be defined by

$$\tau_k^* = \inf\Big\{s \geq \tau_{k-1}^*\colon J^k(s, G_s, I_s, i) = \max_{j\neq i}\big(-C_{i,j} + J^{k-1}(s, G_s, I_s, j)\big)\Big\} \wedge T. \quad (5.4)$$

Let us point out that the inventory process (I_t) is degenerate in the sense that it can be deterministically computed from knowledge of initial inventory level I_0 and switching policy u. In particular, in our model if there are no switches on $[t, \tau)$, I_τ is a linear function of I_t and $(\tau - t)$. For instance, if the initial regime is 'inject' we can re-write equation (5.3) as

$$J^k(t,g,c,-1) = \sup_{\tau\in\mathcal{S}_t} \mathbb{E}\Big[-\int_t^\tau a_{in}\cdot G_s\, ds + \max_{j\neq i}\Big\{-C_{i,j} - b\cdot\int_0^{(\tau-t)}(c + a_{in}\cdot s)\, ds + J^{k-1}\left(\tau, G_\tau, c + a_{in}(\tau - t), j\right)\Big\}\Big| G_t = g\Big],$$

with no mention of I_t at all. This feature will have important implications for the numerical implementation below.

The PDE approach has also been explored for gas storage by deriving the QVI satisfied by the value function [2]. In our case, the basic QVI is

$$\begin{cases} \phi(t,g,c,i) \geqslant \max_{j\neq i}\big(-C_{i,j} + \phi(t,g,c,j)\big), \\ \partial_t\phi(t,g,c,i) + \mathcal{L}_G\phi(t,g,c,i) - a_i\,\partial_I\phi(t,g,c,i) + (a_i\, g - b\cdot c) \leqslant 0. \end{cases} \quad (5.5)$$

The system can now be solved using, for instance, the FD method. Observe that even in the simplest case when (G_t) is one-dimensional, the PDE is two-dimensional in space, making it quite slow.

5.1.1 Backward Recursion For Inventory Level

The difficulty with solving (5.2) by the usual backward dynamic programming algorithm is that at time t we do not know the optimal current inventory I_t^*. Indeed, starting with I_0, I_t^* depends on the past path of (G_t) and the optimal strategy u^* on $[0,t]$. Both are unknown from the point of view of the backward recursion. To overcome this problem, we simply discretize the I-space and keep Longstaff-Schwartz regressions in the G-dimension. In this modified algorithm suppose that we know $J^k(t_2, g_{t_2}^\ell, c, i)$ along the paths $(g^\ell)_{\ell=1}^{N^p}$ for any inventory level c. Now fix the current inventory level $\bar{c}$ and regress all the conditional expectations to derive an estimate for $J^k(t_1, g_{t_1}^\ell, \bar{c}, i)$ in the same way as in Section 4.2:

$$J^k(t_1, g_{t_1}^\ell, \bar{c}, i) = \Big(\hat{E}_{t_1}\big[J^k(t_2, \cdot, \bar{c} + a_i\Delta t, i)\big](g_{t_1}^\ell) + \psi_i(g_{t_1}^\ell, \bar{c})\Delta t\Big) \\ \vee \max_{j\neq i}\big(-C_{i,j} + J^{k-1}(t_1, g_{t_1}^\ell, \bar{c}, j)\big). \quad (5.6)$$

It remains to obtain a full map $c \mapsto J^k(t_1, g_{t_1}^\ell, c, i)$. To do so, vary $\bar{c}$ by using a grid of N^c values $\{c_j \colon c_j = c_0 + j\Delta c\}$ and then interpolate. In principle, for a fixed Δt we can even construct a full grid in the I-dimension and solve the problem exactly. Indeed, if we are in regime i then in one time step Δt, the inventory changes by $a_i \cdot \Delta t$ and so if we take

$$\Delta c \in \{x \colon \exists n_i \in \mathbb{N} \text{ s.t. } x \cdot n_i = a_i \cdot \Delta t \ \forall i\}$$

then inventory adjustments in all regimes result in an integral number of jumps on the inventory grid and no interpolation is needed.

The above approach (with or without interpolation) is very time intensive since we now run a separate regression for each inventory c_j, regime i, number of switches k and time step $m\Delta t$. Moreover, we are no longer able to employ the true Longstaff-Schwartz scheme of approximating the switching times. Because we only have values for J^k at the grid points $(m\Delta t, g_{m\Delta t}^\ell, c_j)$, we must store the conditional expectations

at every time step. Nevertheless, the algorithm is quite robust and is the simplest way of dealing with inventory level parameters.

To maintain numerical efficiency it is desirable to eliminate the fixed discretization in the I-space that resembles the classical lattice schemes with their bad dimensional scaling. Accordingly, we propose the following modification that instead uses *pathwise* inventory levels ($I^\ell_{m\Delta t}$). This allows us to do a joint (G, I)-regression and keep all the features of LS scheme. The idea is to generate for each path and time point $(\ell, m\Delta t)$ a random inventory level $I^\ell_{m\Delta t}$. Then the expected future profit conditional on $I^\ell_{m\Delta t}$ and $g^\ell_{m\Delta t}$ is obtained by a double regression in (5.6) of future values $J((m+1)\Delta t, g^\ell_{(m+1)\Delta t}, I^\ell_{(m+1)\Delta t}, i)$ against the Markovian current state $(g^\ell_{m\Delta t}, I^\ell_{m\Delta t})_{\ell=1}^N$. The extra randomization in (I_t) allows us to reduce computations by leveraging the information from other paths. From another angle, the grids for current $I_{m\Delta t}$ are now random and we use a global regression (rather than local interpolation) to perform the backward recursion. To perform the 2-d regression it is likely that a large number of basis functions is needed (about $10 - 15$ in our experience) which in turn means that a large number of simulations is necessary. Nevertheless, the scheme still handily beats the interpolation method. Furthermore, viewing I_t as a generic history variable we suggest that this idea can be applied to carry out the dynamic programming algorithm for any past-dependent setting.

Example 5.1. For a numerical illustration we consider a facility with a total capacity of 8 MMBtu rented out for one year. The price process is taken from the data of [25],

$$d \log G_t = 17.1(\log 3 - \log G_t)\, dt + 1.33\, dW_t.$$

Observe the very fast mean-reversion of the prices, with a half-life of 20 days. The initial inventory is 4 MMBtu and the terminal condition is $J(T, g, c, i) = -2 \cdot g \cdot (4 - c)^+$. The other parameters in (5.1) are

$$a_{in} = 0.06 \cdot 252, \quad b = 0.1,$$
$$a_{out} = 0.25 \cdot 252, \quad C = 0.25.$$

We solve this toy storage problem using three different solvers: an explicit finite-difference PDE solver discretizing (5.5), a mixed interpolation-regression scheme and a pure 2-d regression scheme. The results are summarized in Table 5.1. We do not have an intuition which method is the most accurate, but it is reassuring to see all three values within 2.5% of each other. Let us also point out the long times required to run each method indicating the computational challenges involved. In this light, the 45% time savings obtained by the (G, I)-regression scheme become crucial from a practical point of view.

Method	Mean	Std. Dev	Time (m)
Explicit FD	6.95	–	55
LS interp	7.11	0.021	47
LS 2-d regression	7.04	0.038	32

Table 5.1: Comparison of numerical results for Example 5.1. The PDE solver used a 400×400 grid and 10000 time steps. The interpolation scheme used 400 time steps, 10000 paths and 80 grid points in the I-dimension. The 2-d regression used 400 time steps and 40000 paths. Standard deviations were obtained by running the Monte Carlo methods 50 times.

Figure 5.1 shows the value function $J(t, g, c, i)$ as a function of current price and inventory for an intermediate time $t = 0.5$ and mode 'store'. Not surprisingly, higher inventory increases the value function since one has the opportunity to simply sell the excess gas on the market. In the G_t-direction we observe a parabolic shape with a minimum around $G_t = 3$. This suggests that deviations of G_t from its mean imply higher future profits, confirming our intuition about storage acting as a straddle.

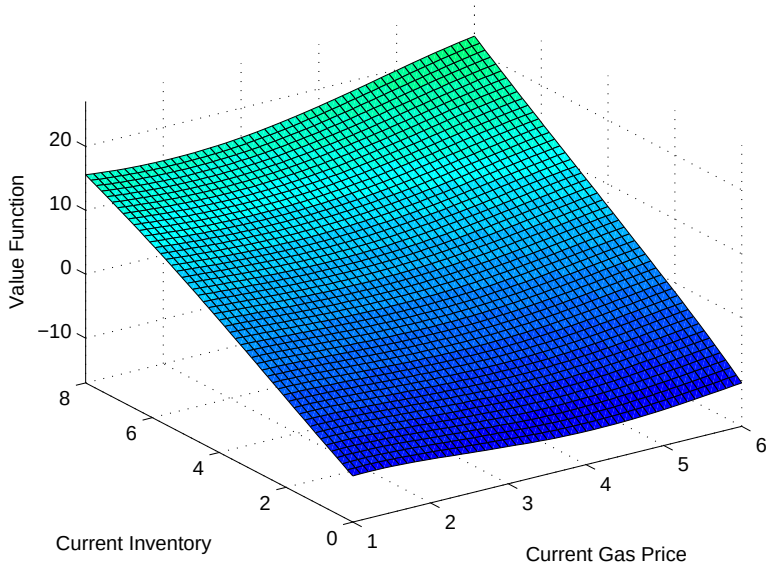


Figure 5.1: Value function surface for Example 5.1 showing $J(0.5, g, c, \text{`}store\text{`}; T = 1)$ as a function of current gas price $G_t = g$ and current inventory $I_t = c$.

5.2 Hedging Supply Guarantees

The model developed in the previous section can be used as a foundation for analyzing the problem of hedging firm supply contracts which are common in the power industry. The basic setup postulates that an energy merchant signs a fixed price contract guaranteeing delivery of power to a utility during a specified period. The merchant owns a power plant that will generate this power. However, there are two major issues. First, the power plant will require gas that must be obtained in advance. Supposing the merchant also owns a gas storage facility, he can begin buying gas early in order to lock in lower prices during the off-season. Second, the final amount of power demanded L_T is *stochastic* and will only be determined at maturity. It will depend on a variety of factors, such as weather, market prices, state of economy, etc. Given the uncertainty of the demand and the flexibility of gas storage, the task of the supplier is to build up a gas inventory to match the terminal demand at the lowest total cost. For simplicity, we assume that the conversion of

gas to power is deterministic so that we can restrict our attention to just the gas market. Mathematically, the agent therefore needs to hedge the amount L_T of gas by buying, selling and storing gas on $[0, T]$ with minimum cost.

To solve the problem we need to specify the joint dynamics of gas prices (G_t) and expected demand load $\tilde{L}_t = \mathbb{E}[L_T | \mathcal{F}_t]$. We assume that the demand can be described by an observed stochastic process (L_t), so that L_T is not a separate random variable but a point on a path of (L_t). One reasonable model is to take the demand to be a mean-reverting process with a time-dependent level. Hence,

$$dL_t = \kappa_L(\theta_t^L - L_t)\, dt + \sigma_L dW_t^2. \tag{5.7}$$

The mean-reverting level θ_t^L can be used to model the seasonal power demand with higher consumption in the peak seasons. We will continue to assume that the gas prices (G_t) follow an exponential OU process, cf. (3.4), correlated with (L_t). Let us also note that under (5.7) the conditional expectation $\tilde{L}_t$ is a martingale with exponentially decaying volatility,

$$d\tilde{L}_t = \sigma_L \mathrm{e}^{-\kappa_L(T-t)} dW_t^2.$$

Regarding the penalty for not having the correct inventory of gas at final date T, we model it as $\bar{C} \cdot G_T(L_T - c_T)^+$. The latter term represents the cost of buying the remaining gas on the market with a premium $\bar{C}$ (e.g. $\bar{C} = 1.5$), and with no option of selling the surplus. Defining the value function as $J(t, g, c, l, i)$, $X_t = (G_t, I_t, L_t, u_t)$ we have a modified gas storage problem given by

$$J(t, g, c, l, i) \triangleq \sup_{u \in \mathcal{U}(t)} \mathbb{E}\Big[\int_t^T \psi_{u_s}(s, G_s, I_s)\, ds - C \int_t^T |du|_s \Big| \; X_t = (g, c, l, i)\Big], \tag{5.8}$$

$$J(T, g, c, l, i) = -\bar{C} \cdot g \cdot (l - c)^+,$$

and with the regimes ψ_i as in (5.1).

Example 5.2. To illustrate the results we re-use the gas storage example 5.1 from the previous section. We assume that

$$dL_t = 2 \cdot (4 \cdot (0.9 + 0.1\cos(4\pi t)) - L_t)\, dt + \sigma_L dW_t^2, \qquad d\langle W^1, W^2\rangle_t = -0.8\, dt,$$

so that the mean expected load is 4 MMBtu or half the facility capacity of 8 MMBtu. The negative correlation between W^1 and W^2 indicates that the demand declines as gas price increases. The fluctuation in the mean level of L_t models the bi-yearly seasonality of demand. Since the average gas price is three dollars per Btu, without the storage facility the agent will need to pay about $12 million plus hedging costs. The agent receives the facility empty, so that with the given injection rate it will take 96 days to inject four MMBtu. The switching costs are set to be pretty high at 0.25 million dollars.

Table 5.2 summarizes the expected costs as we vary the amount of time available and the volatility σ_L of the load. We see that as the amount of time increases, the cost declines and in fact with a horizon of one year the agent has plenty of time to engage in speculative storage trading on top of hedging the demand. We also see that as expected, more volatile load costs more since the agent is unsure till the very end how much would be needed and hence will try to superhedge to avoid the penalty of not having enough. Even if the difference of about 0.1 is small, the relative difference compared to the savings from the storage facility (e.g. about two million over six months) is financially significant.

5.3 Exhaustible Resources

Besides valuing energy tolling agreements, our model is closely related to management of exhaustible resources. In the latter, the firm owns a natural resource, such as a mine or oilfield that it would like to optimally extract. The resource is subject to fluctuating price levels and the firm can stop and restart extraction. On infinite horizon, exhaustible resources have been studied in a classic paper by Brennan and

Maturity in Months	Demand Vol. $\sigma_L = 0.2$	Demand Vol. $\sigma_L = 0.4$	Demand Vol. $\sigma_L = 0.8$
3	12.17	12.19	12.25
6	9.99	10.01	10.08
9	7.53	7.55	7.64
12	5.23	5.25	5.33

Table 5.2: Price of power supply guarantee in millions of USD for Example 5.2. We use the LS interpolation scheme with 8000 paths and a grid of 50 inventory values.

Schwartz [13]. However, for realistic financial planning the horizon should be finite. Indeed, many development licenses have finite lifetimes and the operating company must extract as much as possible by the deadline. For instance, a mining company may have only a 20-year lease on the mine site after which it will have to obtain a new license from the government.

Such resource management is easily convertible into our recursive optimal stopping framework. For simplicity (a standard assumption in the literature), we assume that the mine can be operated in a discrete number of regimes with production rates a_i. Letting I_t denote the current level of resources left (hence I_t is a non-increasing process), the model becomes a direct analogue of gas storage. Again, I_t is a degenerate state variable that depends on past history and whose dynamics can be directly extracted from the choice of managerial policy. Given initial (known) total inventory of c_0 the objective is

$$J(t,x,c,i) = \sup_{\tau,j} \mathbb{E}\Big[\int_t^\tau \psi_i(s,X_s)\,ds + J(\tau,X_\tau,I_\tau,j) - C_{i,j}\Big|\, X_t = x, I_t = c\Big] \tag{5.9}$$

where $I_\tau = c - a_i(\tau - t)$. Resource depletion is modeled with the boundary condition $J(t,x,0,i) \equiv 0$.

Example 5.3. For comparison we take up the example of mine management in [13]. Consider a copper mine with finite inventory of 15 years worth of production. The mine can be either operated at a fixed rate or kept idle. The price process (X_t)

follows a geometric Brownian motion and the payoff rates in annualized millions of dollars (considering taxes and real estate) are

$$\begin{cases} \psi_0(X_t) = -K_0, \\ \psi_1(X_t) = q \cdot (X_t - K) - 0.5 \cdot q \cdot (X_t - K)^+, \\ dX_t = \mu X_t \, dt + \sigma X_t \, dW_t, \\ q = 10, K = 0.5, K_0 = 0.5, \mu = 0.01, \sigma^2 = 0.08. \end{cases}$$

Finally, the switching cost is $C = 0.2$, the discount rate is $r = 4\%$ and the mine may be abandoned, putting a local lower bound of zero on the value of a closed mine.

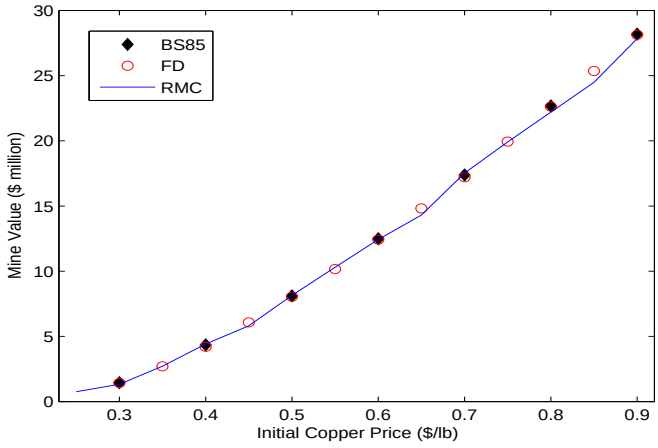


Figure 5.2: Comparison of different pricing methods for the copper mine example of Brennan and Schwartz [13], varying the initial copper price X_0. In [13] only the values $X_0 = 0.1 \cdot k$ were reported, while we computed on a finer grid $X_0 = 0.05 \cdot k$. The PDE solver used a 400×300 grid and 10000 time steps. The mixed regression-interpolation scheme labeled (RMC) used 600 time steps, 10000 paths and 75 gridpoints in the I-direction.

We solve this problem assuming a finite horizon of 99 years that is standard for mine exploitation licenses. We use the mixed regression plus I-interpolation scheme and a FD PDE solver as a double check to make sure we can recover the reference

values. The PDE solver uses finite differences to solve the QVI which for this problem is given by [13, p. 343]

$$\begin{cases} J^1_t + \frac{1}{2}\sigma^2 x^2 J^1_{xx} + \mu x J^1_x - q J^1_q + q(x-K) - 0.5\, q(x-K)^+ - r J^1 = 0, \\ J^0_t + \frac{1}{2}\sigma^2 x^2 J^0_{xx} + \mu x J^0_x - K_0 - r J^0 = 0, \\ J(t,x,q,0) \geqslant 0 \qquad \text{and} \quad |J^1 - J^0| \leqslant C. \end{cases}$$

The results are compared to those reported by Brennan and Schwartz [13] and summarized in Figure 5.2. As we see all three solvers are consistent with each other. Our regression solver has errors of less than 2% everywhere and is an order faster than the PDE method, already showing significant savings for this toy example.

5.4 Incorporating Other Features

As we have admitted all along, the models presented have been simplifications. Let us discuss which features one might want to add for a realistic implementation. First of all, energy prices are observed to be spiky and time-dependent. As mentioned before, upward jumps can be added easily since the Snell envelope will remain continuous and thus the value function stays smooth. In principle, smoothness is not necessary, however the projected conditional expectations will have difficulty approximating non-smooth functions. Similarly, time-dependent coefficients are straightforward to incorporate, as long as we have a method for simulating the forward (X_t) paths. In special cases one can even obtain closed-form results, cf. Section 4.8. Clearly, any agent wishing to price an (X_t)-derivative should already have handy some scheme for such simulations. Switching costs and reward rates are also likely to be time-inhomogenous and $C_{i,j}$ might depend on the current (X_t). Again, as long as the barrier is continuous, this does not affect the qualitative properties of the model.

Some problems may have an abandonment feature. For example, in the exhaustible resources problem, the mine operator had the option of complete abandonment, shutting down all operations and exiting the business. This provided insurance

against negative cash flows if the commodity price dropped really low. Besides [13], an infinite horizon problem with such abandonment option was considered by Zervos [72]. In our framework, the implementation is simple. We can directly impose a constraint of the form $J^k(m\Delta t, x, i) \geqslant Ab(m\Delta t, x, i)$ during dynamic programming, where $Ab(t, x, i)$ represents the total value from abandoning at time t and state (x, i).

In the setting of tolling agreements a crucial feature is outages. Sometimes the plant experiences a malfunction and must be shut down for maintenance. Outages are critical in practice and are responsible for many of the electricity price spikes. They make the operator more risk averse since the benefit of being in an 'on' state is reduced. In general, outages can be broken down into planned and un-planned ones. The planned outages are usually for routine maintenance and can be seen as part of the operational characteristics of the plant. Since they are deterministic, we will ignore them here and concentrate on unplanned emergencies. For simplicity we assume that the latter are completely unpredictable and occur independently with a constant intensity rate λ. Letting $\tilde{T}_k$ represent the random time of the k-th emergency, we have $\mathbb{P}[\tilde{T}_k \geq s + \tilde{T}_{k-1}] = \mathrm{e}^{-\lambda s}$. Hence, the times between outages have exponential distribution, and we obtain the same framework as for exponential maturity randomization, cf. (3.15). The recursive construction (3.7) becomes

$$J^k(t,x,i) = \sup_{\tau \in \mathcal{S}_t} \mathbb{E}\Big[\int_t^\tau \mathrm{e}^{-\lambda(s-t)} \big(\psi_i(s, X_s) + \lambda (J^{k-1}(s, X_s, 0) - C_{i,0}) \big)\, ds \\ + \mathrm{e}^{-\lambda(\tau - t)} \max_{j \neq i} \Big(J^{k-1}(\tau, X_\tau, j) - C_{i,j} \Big) \Big| X_t = x \Big],$$

where the first term represents the probability of an outage and a forced switch to mode 'off' before τ. Numerically, outages are very simple to implement in discrete time. The benefit of no switch at instant t is simply changed to $\mathbb{E}\big[(1 - \lambda\Delta t) J^k(t + \Delta t, X_{t+\Delta t}, i) + \lambda\Delta t (J^{k-1}(t + \Delta t, X_{t+\Delta t}, 0) - C_{i,0}) \big| \mathcal{F}_t \big] + \psi_i(t, X_t)\Delta t$. Related work has been done by Wang [69].

Finally, let us mention some features that are not easily implementable in our framework. First and foremost these include 'memory' properties of operating a plant that destroy the Markov property of (X_t). For instance, in practice the heat rate is a function of the time the plant has been in operation. The plant is less efficient when just started and when running for a long time. Similarly, rapid ramping-up increases the heat rate, meaning that $\psi_i(s, X_s)$ also depends on the last regime used. Secondly, in reality the plant operator is not a price taker and hence the chosen strategy u will in fact influence the price. If this influence has long-term effects our framework breaks down. We are no longer able to simulate (X_t) by itself, since it is now affected by the choice of strategy. Lastly, the already mentioned operational delay features will be hard to implement if we allow the operator to cancel and/or change her mind during the transition period.

5.5 Utility Maximization

So far our optimization problem has been linear in the sense that we were simply optimizing the expected profit. In practice, the agent is likely to be risk-averse and would wish to smooth her realized profits. Here we talk about *total* realized profits on the entire horizon $[0, T]$, which we recall for a given strategy u were labeled as $H(x, i, [0, T]; u)$. Indeed, the agent should care not only about the expected value of $H(x, i, [0, T]; u)$ but also about its riskiness. This suggests we impose a global utility function U to measure the risk-adjusted value of the stochastic cumulative payoff $H(x, i, [0, T]; u)$. We think of U as penalizing variance, so that the agent is more risk-averse in her managing policy. She will be willing to forgo some profitable opportunities when they are highly uncertain in order to smooth out total profit flows. Moreover, this behavior will be a function of her past realized profit.

Our previous setup consisted in optimization of expected profit under the risk-neutral measure. Alternatively, we could reinterpret it as optimization under the

historical measure with a so-called differential or local utility, which can model risk preferences about the local cash flows. Formally, let $\mathbb{P}$ be the historical measure and take $\psi_i(t, X_t) = \bar{U}(t, \phi_i(t, X_t))$ where $\phi_i(t, X_t)$ is the nominal cash flow in regime i and $\bar{U}$ is a concave utility functional. For instance, we might take $\phi_i(t, X_t) = 10(X_t - 10)$ and $\bar{U}(t, x) = \mathrm{e}^{-rt-\gamma x}$, $\gamma > 0$ to model a differential constant relative risk aversion. The advantage of differential utility was linearity. Indeed, the optimal policy on $[t, T]$ was independent of the realized profits on $[0, t]$. This time additivity made the optimal policy a function of the current (X_t) only. With global utility U, this is no longer the case. Let (w_t^u) denote the cumulative wealth process resulting from using strategy u. This will slightly differ from $H(x, i, [0, t]; u)$ since we now allow investment in a risk-free bank account earning interest r. Hence, the (w_t^u) process has deterministic dynamics

$$\begin{cases} dw_t^u = (rw_t + \psi_{u_t}(t, X_t))\, dt & \text{no switch,} \\ w_\tau^u = w_{\tau-} - C(u_{\tau-}, u_\tau) & \text{switch at } \tau. \end{cases} \tag{5.10}$$

If there are no switches on $[t, \tau)$ we obtain

$$w_\tau^u = \Big(w_t^u + \int_t^\tau \mathrm{e}^{-r(s-t)} \psi_{u_s}(s, X_s)\, ds\Big) \mathrm{e}^{r(\tau-t)} = \mathrm{e}^{r(\tau-t)} w_t^u + \int_t^\tau \mathrm{e}^{r(\tau-s)} \psi_{u_s}(s, X_s)\, ds.$$

The new control problem can now be stated as

$$\begin{cases} J(t, x, w, i) = \sup\limits_{u \in \mathcal{U}(t)} \mathbb{E}\Big[U(w_T^u) \Big| X_t = x, w_t = w, u_t = i\Big], \\ J(T, x, w, i) = U(w). \end{cases} \tag{5.11}$$

Applying the dynamic programming principle to (5.11) and letting τ denote the first switching time we obtain

$$\begin{aligned} J(t, x, w, i) &= \sup_{\tau \in \mathcal{S}_t,\, j} \mathbb{E}\big[J(\tau, X_\tau^{t,x}, w_\tau, j) \big| X_t = x, w_t = w, u_t = i\big], \\ \text{with} \qquad w_\tau &= \mathrm{e}^{r(\tau-t)} w + \int_t^\tau \mathrm{e}^{r(\tau-s)} \psi_i(s, X_s^{t,x})\, ds - C_{i,j}. \end{aligned}$$

Superficially, solving (5.11) is much harder because (w_t^u) is backward-looking, in particular depending on the history of chosen policy u. However, thanks to the degenerate dynamics of (w_t^u) we once again can accommodate it numerically in the same manner we incorporate current inventory levels in the gas storage problem. Therefore the whole discussion in Section 5.1.1 is applicable.

The special case where there are no switching costs and $U(w) = \mathrm{e}^{-\gamma w}$ is called risk sensitive control and is related to BSDE's. As shown by El Karoui and Hamadène [33], the general stochastic control problem

$$J(0, x_0, u_0) = \sup_{u \in \mathcal{U}} \mathbb{E}\Big[\exp(-\gamma \int_0^T \psi(s, X_s, u_s)\, ds)\Big|\, X_0 = x\Big] \tag{5.12}$$

can be associated with a quadratic BSDE. Namely, $J(0, x_0, u_0; u) = \exp(Y_0^u)$ where

$$-dY_t^u = \big(Z_t \frac{\mu(X_t)}{\sigma(X_t)} + \psi(t, X_t, u_t) + \frac{\gamma^2}{2}|Z_t|^2\big)\, dt - Z_t\, dW_t, \qquad Y_T = 0.$$

Note that u_t is now allowed to take a continuum of values. Maximizing $J(0, x_0, u_0; u)$ is therefore equivalent to maximizing Y_0^u which in turn can be done by taking the infimum in the drift term above due to the comparison theorem for BSDE's. As a result, one obtains a closed-form expression for the optimal u_t^* and a numerical scheme similar to those discussed in this dissertation could be used to compute the corresponding $Y_0^{u^*}$. It would be interesting to extend this relation further and incorporate switching costs.

5.6 Risk Management and Hedging

The previous section is a first step towards a systematic method of risk management in the presence of managerial flexibility. Consider the trading arm of an energy merchant that owns complex portfolios consisting both of standard financial instruments (European options, futures, etc.) and physical assets with a controllable dispatch policy. Facing this set of derivatives two main issues arise: how to optimally manage the assets given the rest of the portfolio? And how to set up hedges for the assets

using liquid instruments? Both of these fundamental questions hinge on creating a tractable and *consistent* method of assigning risk to a general energy portfolio.

To formalize, envision a portfolio of derivatives on an underlying (X_t). We associate to this portfolio a wealth process $\mathbf{w}$ which counts the cumulative cash flows on $[0,T]$. For example, if the portfolio just contains a European option paying out $\phi(X_T)$ and the initial endowment is zero, then $w_t = \mathbb{1}_{\{T\}}(t) \cdot \phi(X_t)$. Abstractly, a wealth process is any adapted, $(\mathbb{R}_+ \times \Omega, \mathcal{B}_{\mathbb{R}_+} \otimes \mathcal{F}_t)$ measurable, square-integrable stochastic process. Let us stress that $\mathbf{w}$ is cumulative, rather than representing raw cash flows. The financial value of the wealth process $\mathbf{w}$ from a standpoint of view at time t is measured by a utility functional $\rho(\mathbf{w})$. To conform to general notions of risk we would like that ρ be a convex utility functional as defined by Artzner et al. [3]. The canonical example is $\rho(\mathbf{w}) = \inf_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[\mathbf{w}_T]$, and the precise requirements will be given below.

Suppose now that the agent's portfolio contains a 'vanilla' portion with wealth process $\mathbf{w}$, as well as a tolling agreement $Toll$. Then for any fixed strategy u, we now have an associated value of the entire position given by $\rho(\mathbf{w} + T([0,T];u))$, where $T([0,T];u)$ is the corresponding *cash flow process* from the tolling agreement on $[0,T]$. When we wish to emphasize the initial state, we will write $T(x,i,[t,T];u)$. The latter is defined in analogy with the notation for cumulative profit from (2.2),

$$T(x,i,[0,T];u)(t,\omega) = H(x,i,[0,t];u)(\omega). \tag{5.13}$$

By assumption, the risk preferences ρ have been given to us a priori and the definition of 'best' policy u should take them into account. Accordingly we claim that the appropriate control problem to solve now is

$$J(0,x,w,i;\mathbf{w}) = \sup_{u\in\mathcal{U}} \rho\Big(\mathbf{w} + T(x,i,[0,T];u)\Big) \qquad \text{given} \quad X_0 = x, \mathbf{w}_0 = w. \tag{5.14}$$

Thus, $J(0,x,w,i;\mathbf{w})$ measures the total value to be extracted from the tolling contract and the portfolio $\mathbf{w}$ on the time interval $[0,T]$ given initial state. The optimal

switching problem (5.14) also answers the first question we posed—how to value a general portfolio.

To answer the hedging question we must pass to the dynamic paradigm. Thus, we must be able to dynamically assign risk-adjusted value to a portfolio at any intermediate point t. In particular, ρ needs to be replaced by a set of risk measures ρ_t. To achieve this we use the ideas of Cheridito et al. [20] on consistent monetary utility processes. For technical reasons we work in discrete time and with bounded wealth processes that belong to the space $L^\infty_{0,T} \stackrel{\Delta}{=} \{X\colon \sup_{0\leqslant m\Delta t\leqslant T} |X_{m\Delta t}| < \infty\}$ (resp. $L^\infty_{\tau,T}$ where τ is a general $\mathcal{F}^\Delta$-stopping time).

Definition 5. *[20, Definition 3.1] A Mapping $\phi\colon L^\infty_{\tau,T} \to L^\infty(\mathcal{F}_\tau)$ is called a coherent concave monetary utility functional (CCMUF) if it has the following five properties:*

Measurability: $\phi(\mathbb{1}_A X) = \mathbb{1}_A \phi(X)$ *for all* $X \in L^\infty_{\tau,T}$ *and* $A \in \mathcal{F}_\tau$.

Monotonicity: $\phi(X) \leqslant \phi(Y)$ *if* $X_t \leqslant Y_t$ *for all* $t \in [\tau, T]$ *a.s.*

Certainty Invariance: $\phi(X + m\cdot \mathbb{1}_{\tau,T}) = \phi(X) + m$, *for all* $X \in L^\infty_{\tau,T}$ *and* $m \in \mathcal{F}_\tau$.

Concavity: $\phi(\lambda X + (1-\lambda)Y) \geqslant \lambda\phi(X) + (1-\lambda)\phi(Y)$ *for all* $X, Y \in L^\infty_{\tau,T}$ *and* $\lambda \in \mathcal{F}_\tau$ *with* $0 \leqslant \lambda \leqslant 1$.

Coherence: $\phi(\lambda X) = \lambda\phi(X)$ *for all* $X \in L^\infty_{\tau,T}$ *and* $\lambda \geqslant 0, \lambda \in \mathbb{R}_+$.

Definition 6. *[20, Definition 4.1 and 4.3] Suppose that for each $t = m\Delta t$, ϕ_t is a CCMUF on $L^\infty_{t,T}$. For a $\mathcal{F}^\Delta$-stopping time τ, define ϕ_τ by $\phi_\tau(X) = \sum_m \phi_{m\Delta t}(\mathbb{1}_{\tau = m\Delta t}\cdot X)$. The set $\{\phi_\tau\}_{\tau\in\mathcal{S}^\Delta}$ is called a monetary utility process. Such a process $\{\phi_\tau\}$ is called* consistent *if for any $0 \leqslant t \leqslant \tau \leqslant T$,*

$$\phi_t(X) = \phi_t\big(X\mathbb{1}_{[t,\tau)} + \phi_\tau(X)\cdot \mathbb{1}_{[\tau,T]}\big). \tag{5.15}$$

The concept of consistency is supposed to indicate that we can recursively measure the risk-adjusted value of $\mathbf{w}$. Namely, from a starting point t, the total monetary

value of $\mathbf{w}$ is invariant if instead of receiving the random cash flow on $[\tau, T]$, we receive the future value $\rho_\tau(\mathbf{w})$ with certainty at τ and no payments thereafter. Note that we allow $\rho_t(\mathbf{w}) = \rho_t(\mathbf{w}_\bullet)$ to depend on the whole path of $\mathbf{w}$ between the present time t and final T.

We now extend the concept of monetary utility process to value standard derivatives in the presence of a tolling agreement. Let $\mathbf{w}$ be a wealth process corresponding to a vanilla derivative. Given a tolling agreement $Toll$ and recalling the notation $T([t,T];u)$ from (5.13), the risk-adjusted value of $\mathbf{w}$ is defined by

$$\tilde{\rho}_t(\mathbf{w}) \triangleq \operatorname*{ess\,sup}_{u\in\mathcal{U}(t)} \rho_t\big(\mathbf{w} + T([t,T];u)\big) - \operatorname*{ess\,sup}_{u\in\mathcal{U}(t)} \rho_t\big(T([t,T];u)\big). \tag{5.16}$$

This definition is reminiscent of the indifference price concept. The value of $\mathbf{w}$ is the difference between the utility of the portfolio with $\mathbf{w}$, minus the utility of having just the tolling agreement. We observe that $\tilde{\rho}$ has a joint effect of diversification and optimization, so that $\tilde{\rho} \geqslant \rho$. Indeed, letting $T^*([\tau,T];\mathbf{w})$ (resp. $T^*(t, X_t, w_t, u_t, [t,T];\mathbf{w})$) denote the optimal $Toll$ wealth process in the presence of a fixed revenue stream $\mathbf{w}$, and using the superadditivity of ρ we have

$$\begin{aligned}
\tilde{\rho}_t(\mathbf{w}) &= \rho_t(\mathbf{w} + T^*([t,T];\mathbf{w})) - \rho_t(T^*([t,T];0)) \\
&\geqslant \rho_t(\mathbf{w} + T^*([t,T];0)) - \rho_t(T^*([t,T];0)) \\
&\geqslant \big[\rho_t(\mathbf{w}) + \rho_t(T^*([t,T];0))\big] - \rho_t(T^*([t,T];0)) = \rho_t(\mathbf{w}).
\end{aligned}$$

We claim that $\tilde{\rho}$ is the canonical construction of assigning value to $\mathbf{w}$, when already owning $Toll$. As such, $\tilde{\rho}$ answers the second question we posed by providing a method for comparing the merits of different wealth processes $\mathbf{w}$ with respect to $Toll$. Note that the question of using $\tilde{\rho}$ for pricing and the resulting consistency with no-arbitrage is more subtle and we do not address it here. For us $\tilde{\rho}$ is just a method of comparing hedging opportunities that in addition preserves all the 'good' properties of ρ. It is immediate to verify that each $\tilde{\rho}_t$ is a CCMUF. Moreover, as the

next theorem demonstrates, $\tilde{\rho}$ inherits the time-consistency of ρ. Intuitively, this is due to Bellman's principle that makes optimization consistent in time.

Theorem 9. *Suppose that ρ is a consistent risk measure. Then $\tilde{\rho}$ is consistent.*

Proof. We need to verify that (5.15) holds for $\tilde{\rho}$. We begin by computing

$$\begin{aligned}\tilde{\rho}_\tau(\mathbf{w}) &= \operatorname*{ess\,sup}_{u\in\mathcal{U}(\tau)} \rho_\tau(\mathbf{w}\mathbb{1}_{[\tau,T]} + T(\tau, X_\tau, u_\tau, [\tau,T]; u)) - \operatorname*{ess\,sup}_{u\in\mathcal{U}(\tau)} \rho_\tau(T(\tau, X_\tau, u_\tau, [\tau,T]; u)) \\ &= \rho_\tau(T^*([\tau,T]; \mathbf{w}\mathbb{1}_{[\tau,T]})) - \rho_\tau(T^*([\tau,T]; 0)). \end{aligned} \tag{5.17}$$

Next we observe that by the dynamic programming principle, the additivity of cumulative profits, and consistency of ρ,

$$\begin{aligned}\operatorname*{ess\,sup}_{u\in\mathcal{U}(t,T)} \rho_t(T(t,x,i,[t,T];u)) &= \operatorname*{ess\,sup}_{u\in\mathcal{U}(t,\tau),\, v\in\mathcal{U}(\tau,T)} \rho_t\Big(T(t,x,i,[t,\tau];u)\mathbb{1}_{[t,\tau)} \qquad (5.18)\\ &\qquad + \big[T(\tau, X_\tau, u_\tau, [\tau,T]; v) + H(x,i,[t,\tau];u)\big]\mathbb{1}_{[\tau,T]}\Big) \\ &= \operatorname*{ess\,sup}_{u\in\mathcal{U}(t,\tau)} \rho_t\Big(T(t,x,i,[t,\tau];u)\mathbb{1}_{[t,\tau)} \qquad (5.19)\\ &\qquad + \operatorname*{ess\,sup}_{v\in\mathcal{U}(\tau,T)} \rho_\tau\big(T(\tau, X_\tau, u_\tau, [\tau,T]; v) + H(x,i,[t,\tau];u)\big)\mathbb{1}_{[\tau,T]}\Big).\end{aligned}$$

The last expression shows that we can split the optimization step into recursive optimization over $[\tau, T]$ and then over $[t, \tau)$. We are now ready to compute

$$\begin{aligned}&\tilde{\rho}_t\Big(\mathbf{w}\mathbb{1}_{[t,\tau)} + \tilde{\rho}_\tau(\mathbf{w})\mathbb{1}_{[\tau,T]}\Big) = \\ &\qquad \operatorname*{ess\,sup}_{u\in\mathcal{U}(t,T)} \rho_t\Big(\mathbf{w}\mathbb{1}_{[t,\tau)} + \tilde{\rho}_\tau(\mathbf{w})\mathbb{1}_{[\tau,T]} + T([t,T];u)\Big) - \rho_t\Big(T^*([t,T];0)\Big).\end{aligned}$$

It remains to manipulate RHS_1, the first term on the right hand side:

$$\begin{aligned}
RHS_1 &= \operatorname*{ess\,sup}_{u\in\mathcal{U}(t,\tau)} \rho_t\Big(\mathbf{w}\mathbb{1}_{[t,\tau)} + T([t,\tau];u)\mathbb{1}_{[t,\tau)} \\
&\qquad + \big[\operatorname*{ess\,sup}_{v\in\mathcal{U}(\tau,T)} \rho_\tau\big(T(\tau,X_\tau,u_\tau,[\tau,T];v) + H(x,i,[t,\tau];u) + \tilde{\rho}_\tau(\mathbf{w})\mathbb{1}_{[\tau,T]}\big)\big]\mathbb{1}_{[\tau,T]}\Big) \\
&= \operatorname*{ess\,sup}_{u\in\mathcal{U}(t,\tau)} \rho_t\Big([\mathbf{w}+T([t,\tau];u)]\mathbb{1}_{[t,\tau)} + \big[\rho_\tau(T^*([\tau,T];0)) + H(x,i,[t,\tau];u) \\
&\qquad + \Big\{\rho_\tau(T^*([\tau,T];\mathbf{w}\mathbb{1}_{[\tau,T]})) - \rho_\tau(T^*([\tau,T];0))\Big\}\big]\mathbb{1}_{[\tau,T]}\Big) \\
&= \operatorname*{ess\,sup}_{u\in\mathcal{U}(t,\tau)} \rho_t\Big([\mathbf{w}+T([t,\tau];u)]\mathbb{1}_{[t,\tau)} + \rho_\tau\big(T^*([\tau,T];\mathbf{w}\mathbb{1}_{[\tau,T]}) + H(x,i,[t,\tau];u)\big)\mathbb{1}_{[\tau,T]}\Big) \\
&= \operatorname*{ess\,sup}_{u\in\mathcal{U}(t,T)} \rho_t(\mathbf{w}+T([t,T];u)).
\end{aligned}$$

The first line above applies (5.18), the second line uses (5.19) and (5.17), which due to the shift invariance of ρ_τ leads to terms cancelling each other on the third line, and the last line finally applies (5.19) in reverse. Summarizing,

$$\tilde{\rho}_t\Big(\mathbf{w}\mathbb{1}_{[t,\tau)} + \tilde{\rho}_\tau(\mathbf{w})\mathbb{1}_{[\tau,T]}\Big) = \rho_t(T^*([t,T];\mathbf{w})) - \rho_t(T^*([t,T];0)) = \tilde{\rho}_t(\mathbf{w}),$$

which is precisely the consistency condition. □

The framework can now be used to analyze static hedging by the agent. A European option paying out $\phi(X_T)$ and costing p at time t is worthwhile to purchase given initial holdings $\mathbf{w}$ as long as,

$$\operatorname*{ess\,sup}_{u\in\mathcal{U}(t)} \rho_t\Big(\phi(X_T)\cdot\mathbb{1}_{t=T} + \mathbf{w} + T([t,T];u) - p\Big) > \operatorname*{ess\,sup}_{u\in\mathcal{U}(t)} \rho_t\Big(\mathbf{w}+T([t,T];u)\Big),$$

$$\text{or simply} \qquad \tilde{\rho}_t\big(\phi(X_T)\cdot\mathbb{1}_{t=T}+\mathbf{w}\big) > \tilde{\rho}_t(\mathbf{w}) + p.$$

Other claims of American, Asian, etc. type can be similarly considered. We can analyze more complicated cases as well, for instance where the derivative ϕ is written on a different asset ($X_{1\,t}$) than the underlying in the toll ($X_{2\,t}$). This is the *incomplete markets* situation one encounters in practice when the *Toll* underlying ($X_{2\,t}$) are local power and gas prices (e.g. Chicago city gate), while the financial derivatives

will be written on the liquidly traded NYMEX contracts. From the point of view of optimization, there is no difficulty in extending the model to $X_t = (X_{1\,t}, X_{2\,t})$, and $\tilde{\rho}$ then allows us to price together traded options and the illiquid over-the-counter tolling agreements. In this sense our method is a stochastic control analogue of the framework in McCardle and Smith [58], who constructed a general static paradigm for valuing tradeable and non-tradeable assets. Finally, dynamic hedging should be possible once one appropriately defines continuous-time monetary utility functionals.

On a computational level, tractability depends on the structure of the risk measure ρ. For example, the simplest choice $\rho_t(\mathbf{w}) = \mathbb{E}[\mathbf{w}_T | \mathcal{F}_t]$ corresponds to the linear profit maximization model that we started with in this thesis, and the consistency of ρ reduces to the trivial tower property of conditional expectations. A more interesting case is $\rho_t(\mathbf{w}) = \operatorname{ess\,inf}_{\tau \in \mathcal{S}_t} \mathbb{E}[\mathbf{w}_\tau | \mathcal{F}_t]$ corresponding to the worst-stopping case. Because of the evident relationship between ρ and Snell envelopes our numerical scheme can also be adapted to this situation. We leave further research in this direction to future work.

5.7 Conclusion and Future Research

This dissertation investigated the numerical solution of optimal switching problems that arise in energy markets. Our scheme is based on converting the problem into recursive optimal stopping and then computing the Snell envelopes by backward recursion. We proposed a new scheme based on Monte Carlo regressions that are run simultaneously for all the optimal stopping problems. While our convergence analysis is incomplete, we hope that the breadth of examples listed is compelling as to the merit of this approach. We believe that when it comes to practical implementation our method is more robust and versatile than any that appeared so far in the academic literature. Moreover, it is superior to the currently practiced approach of spread option strips that ignores the operational constraints of tolling agreements.

As the last chapter has demonstrated, the techniques described are applicable to many other energy-related problems. In particular, the gas storage and power supply guarantees are important practical challenges on which our method can shed new light. Full analysis of these two settings will appear in our forthcoming paper [56]. Before concluding, let us mention a couple more possible avenues of investigation. Partially observed models is a common feature of real-life applications. In the simplest setting (X_t) has unobserved components that affect the dynamics but not the profit. An example is the class of convenience yield models that have been popular for energy modeling [16]. In general, stochastic control with partial observations is very challenging numerically since the full state space is infinite dimensional. However several simplifications seem to present themselves. First, if the model is Gaussian, a Kalman filter can be applied, so that the distribution of the unobserved factor is fully described by its conditional mean and variance. Hence, the effective dimension of the problem is just $d + 2$. Then the regression scheme can be applied in the larger dimension after one simulates the observed part $(X_{1\,t})$ and filters the unobserved $(X_{2\,t})$. In the more general case, one could attempt to use particle filters [55]. This means approximating the conditional distribution of $X_{2\,t}$ by a cloud of N^f particles. The effective dimension is reduced to $d + N^f$, but this is still intractable for computing conditional expectations. However, one could attempt to find summary statistics to roughly describe the distribution of the particles. This dimension reduction idea relates back to Barraquand and Martineau [7] who wrote the original paper on solving high-dimensional optimal stopping problems by simulation. It seems that the same approach should work for optimal switching with partial observations. For example, one could parametrize via the moments of the empirical particle distribution.

Stochastic games is another interesting application. Instead of having a single controller, consider the situation where there are two players with competing inter-

ests. As long as simultaneous switches are not allowed and there is a clear rule on the precedence of switching decisions, the problem should fall into the realm of optimal switching we have analyzed here. In general, stochastic games are naturally related to doubly-reflected BSDE's, see e.g. [36]. Possible finance applications might include exotic convertible bonds, compound game options and competition games between firms.

Returning to optimal switching itself, several open problems remain that we would like to settle. A better grip on the regression and Monte Carlo error propagation is needed to understand the convergence properties of the LS scheme. Also, it would be very useful to have tractable conditions regarding existence of connected switching sets, extending Theorem 5. The method of excessive majorants should be further investigated as it might provide new closed form solutions (or at least recursive formulae) for the value functions. On a more fundamental level, deeper understanding of the connection between quasi-variational inequalities and reflected BSDE's that we briefly touched upon in Section 3.7 has the potential to provide new probabilistic tools for numerically solving the classical Stefan problem.

References

[1] A. Abel and J. Eberly. An exact solution for the investment and value of a company facing uncertainty, adjustment costs, and irreversibility. *J. Econom. Dynam. Control*, 21:831–852, 1997.

[2] H. Ahn, A. Danilova, and G. Swindle. Storing arb. *Wilmott*, 1, 2002.

[3] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Math. Finance*, 9(3):203–228, 1999.

[4] V. Bally and G. Pagès. Error analysis of the optimal quantization algorithm for obstacle problems. *Stochastic Process. Appl.*, 106(1):1–40, 2003.

[5] V. Bally, G. Pagès, and J. Printems. A quantization tree method for pricing and hedging multidimensional American options. *Mathematical Finance*, 15(1):119–168, 2005.

[6] A. Bar-Ilan, A. Sulem, and A. Zanello. Time-to-build and capacity choice. *J. Econom. Dynam. Control*, 26(1):69–98, 2002.

[7] J. Barraquand and D. Martineau. Numerical valuation of high-dimensional multivariate American securities. *Journal of Financial and Quantitative Analysis*, 30(3):383–405, 1995.

[8] A. Bensoussan and J.-L. Lions. *Impulse Control and Quasi-Variational Inequalities*. Gauthier-Villars, Paris, 1984.

[9] E. Boguslavskaya. On optimization of dividend flow for a company with non-negative liquidation value. 2003. working paper.

[10] B. Bouchard, N. El-Karoui, and N. Touzi. Maturity randomization. 2004. preprint.

[11] B. Bouchard and N. Touzi. Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Process. Appl.*, 111(2):175–206, 2004.

[12] K. A. Brekke and B. Øksendal. Optimal switching in an economic activity under uncertainty. *SIAM J. Control Optim.*, 32(4):1021–1036, 1994.

[13] M. Brennan and E. Schwartz. Evaluating natural resource investments. *Journal of Business*, 58:135–157, 1985.

[14] A. Cadenillas, R. J. Elliott, and L. A. Léger. On the pricing of American options when the asset follows a mean-reverting process, 2002. Working Paper.

[15] R. Carmona and V. Durrleman. Pricing and hedging spread options. *SIAM Rev.*, 45(4):627–685, 2003.

[16] R. Carmona and M. Ludkovski. Spot convenience yield models for the energy markets. In *Mathematics of finance*, volume 351 of *Contemp. Math.*, pages 65–79. Amer. Math. Soc., Providence, RI, 2004.

[17] R. Carmona and N. Touzi. Optimal multiple stopping and valuation of swing options, 2003. Preprint.

[18] P. Carr. Randomization and the American put. *Review of Financial Studies*, 11(3):597–626, 1998.

[19] J. F. Carrière. Valuation of the early-exercise price for options using simulations and nonparametric regression. *Insurance : Mathematics and Economics*, 19:19–30, 1996.

[20] P. Cheridito, F. Delbaen, and M. Kupper. Dynamic monetary risk measures for bounded discrete-time processes.

[21] D. Chevance. Numerical methods for backward stochastic differential equations. In L. Rogers and D. Talay, editors, *Numerical methods in finance*, pages 232–244. Cambridge University Press, 1997.

[22] E. Clément, D. Lamberton, and P. Protter. An analysis of a least squares regression algorithm for American option pricing. *Finance and Stochastics*, 6:449–471, 2002.

[23] S. Dayanik and M. Egami. Solving impulse control problems via optimal stopping for one-dimensional diffusions, 2004. working paper.

[24] S. Dayanik and I. Karatzas. On the optimal stopping problem for one-dimensional diffusions. *Stochastic Process. Appl.*, 107(2):173–212, 2003.

[25] C. de Jong and K. Walet. To store or not to store. Technical report, Lacima Research Forum, http://www.eprm.com, 2003.

[26] A. Dixit. Entry and exit decisions under uncertainty. *Journal of Political Economy*, 97(3):620–638, June 1989.

[27] J. Duan and S. Pliska. Option valuation with co-integrated asset prices. *Journal of Economic Dynamics and Control*, 28:727–754, 2004.

[28] K. Duckworth and M. Zervos. A model for investment decisions with switching costs. *Ann. Appl. Probab.*, 11(1):239–260, 2001.

[29] P. Dupuis and H. Wang. On the convergence from discrete to continuous time in an optimal stopping problem. *Annals of Applied Probability*, to appear, 2004.

[30] E. B. Dynkin. Optimal choice of the stopping moment of a Markov process. *Dokl. Akad. Nauk SSSR*, 150:238–240, 1963.

[31] E. B. Dynkin. *Markov processes. Vols. I, II*, volume 122 of *Die Grundlehren der Mathematischen Wissenschaften, Bände 121*. Academic Press Inc., Publishers, New York, 1965.

[32] N. El-Karoui. *Ecole d'été de Saint-Flour*, volume 876 of *Lecture Notes in Mathematics*, chapter Les aspects probabilistes du contrôle stochastique, pages 73–258. Springer, Berlin, 1981.

[33] N. El-Karoui and S. Hamadène. BSDEs and risk-sensitive control, zero-sum and nonzero-sum game problems of stochastic functional differential equations. *Stochastic Process. Appl.*, 107(1):145–169, 2003.

[34] N. El Karoui, M. Jeanblanc-Picqué, and S. E. Shreve. Robustness of the Black and Scholes formula. *Math. Finance*, 8(2):93–126, 1998.

[35] N. El-Karoui, C. Kapoudjian, E. Pardoux, S. Peng, and M. Quenez. Reflected solutions of backward SDE's and related obstacle problems for PDE's. *Ann. Probab.*, 25:702–737, 1997.

[36] N. El Karoui, S. Peng, and M. C. Quenez. Backward stochastic differential equations in finance. *Math. Finance*, 7(1):1–71, 1997.

[37] I. Elsanosi, B. Øksendal, and A. Sulem. Some solvable stochastic control problems with delay. *Stochastics Stochastics Rep.*, 71(1-2):69–89, 2000.

[38] A. Eydeland and K. Wolyniec. *Energy and Power Risk Management: New Developments in Modeling, Pricing and Hedging*. John Wiley&Sons, Hoboken, NJ, 2003.

[39] S. Finch. Ornstein Uhlenbeck process, 2004. Preprint.

[40] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*, volume 25 of *Applications of Mathematics*. Springer-Verlag, New York, 1993.

[41] P. Glasserman and B. Yu. Number of paths versus number of basis functions in American option pricing. *Annals of Applied Probability*, 14(4):2090–2119, 2004.

[42] E. Gobet, J. Lemor, and X. Warin. A regression-based Monte-Carlo method to solve backward stochastic differential equations. *Annals of Applied Probability*, 2005. To appear.

[43] X. Guo and H. Pham. Optimal partially reversible investment with entry decision and general production function. *CNRS Preprint n^o* 904, 2004.

[44] S. Hamadène and M. Jeanblanc. On the starting and stopping problem: application in reversible investments. 2004. preprint.

[45] S. Hamadène and J.-P. Lepeltier. Reflected BSDEs and mixed game problem. *Stochastic Process. Appl.*, 85(2):177–188, 2000.

[46] M. Haugh and L. Kogan. Pricing American options: A duality approach. *Oper. Res.*, 52(2):258–270, March-April 2004.

[47] M. Jeanblanc. Impulse control method and the exchange rate. *Mathematical Finance*, 3:161–178, 1993.

[48] M. Jeanblanc and A. Shiryaev. Optimization of the flow of dividends (in russian). *Russian Math Surveys*, 50:257–277, 1995.

[49] I. Karatzas and S. Shreve. *Methods of Mathematical Finance*. Applications of Mathematics, Vol 39. Springer-Verlag, New York, 1998.

[50] T. Kluge. Pricing options in electricity markets, 2004. preprint.

[51] R. Korn. Portfolio optimization with strictly positive transaction costs and impulse control. *Finance and Stochastics*, 2:85–114, 1998.

[52] H. J. Kushner and P. Dupuis. *Numerical methods for stochastic control problems in continuous time*, volume 24 of *Applications of Mathematics*. Springer-Verlag, New York, second edition, 2001.

[53] J.-P. Lepeltier and B. Marchal. Théorie générale du contrôle impulsionnel markovien. *SIAM J. Control Optim.*, 22(4):645–665, 1984.

[54] F. Longstaff and E. Schwartz. Valuing American options by simulations: a simple least squares approach. *The Review of Financial Studies*, 14:113–148, 2001.

[55] M. Ludkovski and R. Carmona. Commodity forwards with partial observations and exponential utility. 2004. Submitted.

[56] M. Ludkovski and R. Carmona. An impulse control approach to gas storage and power supply guarantees. 2005. in preparation.

[57] R. Lumley and M. Zervos. A model for investments in the natural resource industry with switching costs. *Math. Oper. Res.*, 26(4):637–653, 2001.

[58] K. F. McCardle and J. E. Smith. Valuing oil properties: Integrating option pricing and decision analysis approaches. *Operations Research*, 46:198–217, 1998.

[59] M. Mrad, N. Touzi, and A. Zeghal. Monte Carlo estimation of a joint density using Malliavin calculus and applications to Amerian options. Technical report, CREST, France, 2004. preprint.

[60] A. Øksendal. Irreversible investment problems. *Finance Stoch.*, 4(2):223–250, 2000.

[61] B. Øksendal. *Stochastic differential equations*. Springer-Verlag, Berlin, fifth edition, 1998. An introduction with applications.

[62] B. Øksendal and A. Sulem. *Applied stochastic control of jump diffusions*. Springer-Verlag, Berlin, 2005.

[63] G. Pagès. Grilles de quantification optimales. Technical report, World Wide Web resource, http://www.proba.jussieu.fr/pageperso/quant3.html, 2004.

[64] P. Pellizzari and C. Pizzi. Monte Carlo pricing of American options using non-parametric regression. *Rendiconti per gli Studi Economici Quantitativi*, pages 75–91, 2002.

[65] P. Protter. *Stochastic integration and differential equations*, volume 21 of *Applications of Mathematics*. Springer-Verlag, Berlin, 2004. 2nd Edition.

[66] A. N. Shiryayev. *Optimal stopping rules*. Springer-Verlag, New York, 1978. Translated from the Russian by A. B. Aries, Applications of Mathematics, Vol. 8.

[67] J. N. Tsitsiklis and B. Van Roy. Regression methods for pricing complex American-style options. *IEEE Transactions on Neural Networks*, 12(4):694–703, July 2001.

[68] S. Villeneuve and A. Zanette. Parabolic ADI methods for pricing American options on two stocks. *Math. Oper. Res.*, 27(1):121–149, 2002.

[69] H. Wang. Optimal stopping with forced exits. *Math. Oper. Res.*, to appear, 2005.

[70] P. Wilmott, S. Howison, and J. Dewynne. *The mathematics of financial derivatives*. Cambridge University Press, Cambridge, 1995. A student introduction.

[71] A. Yushkevich. Optimal switching problem for countable Markov chains: average reward criterion. *Math. Methods Oper. Res.*, 53(1):1–24, 2001.

[72] M. Zervos. A problem of sequential entry and exit decisions combined with discretionary stopping. *SIAM J. Control Optim.*, 42(2):397–421, 2003.