

# Pricing Asset Scheduling Flexibility using Optimal Switching

René Carmona

Department of Operations Research and Financial Engineering, Bendheim Center for Finance, Princeton University,  
Princeton, NJ 08544 USA, rcarmona@princeton.edu,

Michael Ludkovski

Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA, mludkov@umich.edu,  
www.umich.edu/mludkov

We study the financial engineering aspects of operational flexibility of energy assets. The current practice relies on a representation that uses strips of European spark-spread options, ignoring the operational constraints. Instead, we propose a new approach based on a stochastic impulse control framework. The model reduces to a cascade of optimal stopping problems and directly demonstrates that the optimal dispatch policies can be described with the aid of ‘switching boundaries’, similar to the free boundaries of standard American options. Our main contribution is a new method of numerical solution relying on Monte Carlo regressions. The scheme uses dynamic programming to efficiently approximate the optimal dispatch policy along the simulated paths. Convergence analysis is carried out and results are illustrated with a variety of concrete examples. We benchmark and compare our scheme to alternative numerical methods.

*Key words:* optimal switching, Monte Carlo, operational flexibility, impulse control, Snell envelope

*Subject classifications:* Decision Analysis; Dynamic programming; Applications; Industries: Energy;

*Area of review:* FE

*MSC2000 subject classification:* Primary: 93E20, 65C05; Secondary: 60G40, 60H35

*OR/MS subject classification:* Primary: Dynamic programming/Optimal Control/Applications; Secondary:

Natural Resources/Energy

*History:*

---

## 1. Introduction.

Pricing of operational flexibility is one of the fundamental problems encountered in exotic energy derivatives. To reduce the large liquidity risk present in these markets, trading firms are increasingly seeking control of generating assets like power plants, pipelines, storage facilities, oil refineries, etc. The ownership is transferred by signing temporary lease agreements, and it is of interest to efficiently value such *tolling contracts* given the volatile commodity prices. The complexity of the problem arises due to multiple operational constraints imposed on the manager coupled with a multi-dimensional state process and intricate time-dependence. Features such as price spikes, operational delay, seasonality and market power must be incorporated while maintaining transparency and tractability.

Traditionally, the operational flexibility was represented as a sequence of right-to-run options owned by the manager. This allows to import the extensive intuition built-up from equity and fixed income trading as well as apply extremely fast valuation techniques. However, critically the method is unsuitable for dealing with constraints mentioned above. Switching costs associated with changing the regime of the asset, upper limits on the number of restarts or time-dependency due to contract expiration are all very difficult to handle in this setting. To rectify these shortcomings, we propose to apply a continuous-time stochastic control framework that combines the intuition of American option pricing with a more structural approach. The controls correspond to the entire operational strategy adopted by the agent throughout the contract period. This dynamic *optimal*

*switching* setting is the only way to fully capture the interplay between limited flexibility and uncertainty. The origins of our method go back to Brennan and Schwartz (1985), but only recently it has been rigorously studied.

The existing mathematical finance literature (Dixit 1989, Brekke and Øksendal 1994, Zervos 2003, Guo and Pham 2005, Pham and Ly Vath 2005) related to the optimal switching problem has focused on obtaining closed-form solutions. To do so requires making simplifying assumptions, in particular requiring the state process to be of a simple functional form (e.g. a one-dimensional Geometric Brownian Motion with constant coefficients) and assuming an infinite horizon. In contrast, anticipating real-life applications we do not impose any parametric forms on our model. We work with general Markov multi-dimensional stochastic price processes and consider finite horizon problems. By concentrating on contracts with fixed expiry dates we stay closer to reality, but we face a much more difficult problem. In particular, time decay has to explicitly enter into all the calculations. Moreover, we allow complex operational constraints, including switching delay, general switching costs and forced shutdowns. In addition, our method can easily incorporate price impact which is often observed in the thinly traded energy markets.

The main contribution of our work is a robust numerical scheme for valuing operational flexibility on a finite horizon. The algorithm is simulation-based and, as a result, is easily scalable and flexible to extensions. We give a detailed numerical analysis and perform exhaustive benchmarking to compare several possible numerical implementations. Besides a numerical scheme, we also prove that the optimal switching decisions can be fully described with the aid of switching boundaries. Accordingly, an optimal scheduling strategy has the usual intuitive form of “choose the best alternative given the present expectations”.

The optimal scheduling problem has been topical for practitioners since the beginning of energy market de-regulation in 1970s. Nowadays, nearly every large utility and energy merchant company has a special scheduling desk, where quantitative analysts decide the asset dispatch policy on a daily basis given market conditions. Our approach maintains the currently popular perspective of compound optionality and provides a more flexible and mathematically rigorous formulation of the operational constraints. Coupled with its numerical efficiency, it should have potential for immediate applications in practice.

The organization of this paper is as follows. Section 2 introduces the financial engineering problem we study and carefully formulates the precise mathematical framework. In Section 3 we show that the model can be efficiently reduced to a cascade of recursive optimal stopping problems. From there, we develop in Section 4 a new numerical algorithm based on Monte Carlo simulations and Dynamic Programming. To provide empirical evidence we highlight in Section 5 our methodology in a variety of numerical and qualitative examples. Finally, in Section 6 we discuss additional cases of interest and avenues for further research.

*After this work was completed, we learned from S. Tompaidis of the existence of a similar attempt Deng and Xia (2005) to price tolling agreements. While addressing similar issues, Deng and Xia (2005) seems to be restricted to the Tsitsiklis and van Roy form (cf. (32)) of the regression solution of the dynamic programming equations. Moreover, it does not provide the kind of mathematical analysis presented here (with identification of the continuation and exercise regions), and it does not give the rigorous convergence analysis of the numerical algorithms as we do.*

## 2. Problem Setup.

### 2.1. Tolling Agreements.

Despite their recent growth, energy markets remain relatively illiquid and inefficient. As a result access to the physical ‘equipment’ has extra benefits that cannot be attained otherwise. Many commodity contracts require physical settlement which necessitates actual ownership of an asset.

Moreover, ability to schedule the assets can influence the trading environment and lead to pricing power. Accordingly, financial players in the energy markets have been showing increasing interest in owning energy assets. To circumvent the extremely capital intensive construction and maintenance aspect of the business, the idea of a *tolling agreement* was invented. A tolling agreement temporarily transfers the scheduling flexibility of the asset in return for a fixed payment. For example, a typical tolling agreement gives control of a power plant for a period of one year. From a financial point of view, a tolling agreement is a compound option on the *spread* between the input (natural gas, crude oil, etc.) and the output (electricity, gasoline, etc.) commodity.

To be concrete, let us concentrate on the optimal behavior of a renter that has leased 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. To do so, she will optimize the dispatching policy, i.e. dynamically decide when the plant is running and when it is offline, as time and market conditions evolve. Our aim is to price the flexibility embedded in selecting this dispatching policy. In order to focus on the various operational constraints and the basic structure of the problem, we do not discuss any hedging of the operations; thus the pricing is done in the simplified manner of finding the optimal expected profit under the given pricing measure. In practice, the operational risk involved is non-tradeable, however risk management can still be undertaken using closely-related traded contracts; see Ludkovski (2006) for more details.

## 2.2. Historical Perspective.

Let us briefly summarize the methods that have been used in practice by the financial engineers who are in charge of plant scheduling and who are also responsible for valuing potential tolling agreements and other revenue-sharing contracts. In the early days the standard approach was based on the classical Net Present Value (NPV) theory, also known as discounted cash-flow analysis. The value of the asset was estimated based on projections of future prices and proper weighing and discounting of possible cases. This essentially eliminated uncertainty, and opportunities of dynamically responding to prices were ignored.

In late 1980s, the Markov Decision Processes (MDP) approach became popular. The MDP's are essentially tree-based versions of the stochastic control formulation below. The MDP lattice uses Dynamic Programming to solve the switching problem through backward recursion up the tree. An example of explicit solution for a simple switching problem can be found in Yushkevich (2001). 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 strip of spark-spread options approach (Eydeland and Wolyniec 2003). The payoff from the power plant is represented as a collection of European options that pay the maximum value to be obtained during each decision period. Each such option is of the spark-spread variety which have been intensively studied in classical option pricing (Carmona and Durrleman 2003). As a result, the general intuition that practitioners have built for dealing with vanilla options can be directly transferred. Moreover, in many cases there are efficient approximations for pricing spread options that are very fast to compute and provide good bounds not just on the price but also on the Greek sensitivities. Nevertheless, the method suffers from two major shortcomings. First, it is inherently unsuitable for incorporating *dynamic* operational constraints. For example, there is no natural way of including switching costs. Second, the strip-of-options approach eliminates the time decay, making the resulting strategy completely stationary in time despite the finite horizon of the contract.

Given these difficulties and the growing need for aggressive finetuning of dispatch policies (as the markets become increasingly competitive), we believe it is high time for a next generation of optimal switching models that make full use of the stochastic control formulation and corresponding

numerical methods. Moreover, the new models should be robust so as to handle a variety of operational constraints, a large number of operational regimes, multi-dimensional state processes and varied price dynamics. We hope this paper is a modest step in that direction.

### 2.3. Representative Case.

Let us consider for the sake of illustration, an operator in charge of a peaking combined cycle gas turbine (CCGT) power plant. This is a medium size power plant characterized by short ramping times. The typical capacity is 40 – 200 megawatt-hours (MWh) and it takes 2 – 4 hours to bring the plant online. 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  $\overline{HR}$  is called the *heat rate* and is quoted in millions of British thermal units (MMBtu) of gas needed to produce one MWh (in Europe heat rates are quoted in megawatt-hours per giga-joule MWh/GJ). Thus, the higher the heat rate the less efficient the conversion. Typical heat rates are in the range 7 – 12 MWh/MMBtu. To formulate the problem in financial terms, 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 or day-ahead prices if all the commitments are made on a 24-hour basis. Moreover, we are completely agnostic about the actual dynamics of the  $(P_t)$ - and  $(G_t)$ -processes.

The revenue rate from running the plant is 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.

We suppose that besides running the plant at full capacity, 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 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  $m$  we associate a corresponding *marginal* heat rate  $HR_m$ , production rate  $Cap_m$  and operating expenses  $K_m$ . The relationship between production rates  $Cap_m$  and heat rates  $HR_m$  is likely to be non-linear as increasing losses from heat-dissipation typically cause dis-economies of scale. The rate of payoff in regime  $m$  is then given by

$$\psi_m(P_t, G_t) \triangleq Cap_m \cdot (P_t - HR_m \cdot G_t - K_m). \quad (1)$$

In general, the operating regime  $m$  can affect the dynamics of  $(P_t, G_t)$  through *price impact*. For example, if the firm is a major player on the electricity market, when it runs its turbines we expect the electricity price  $P_t$  to decrease due to increased supply. This effect can be both deterministic (e.g. the price decreases by 5%) or random; in the latter case we make the simplifying assumption that it can be described as changing some parameters of the equation governing the evolution of  $(P_t, G_t)$ , cf. (8). To establish a more abstract notation, we will denote by  $X_t$  the driving Markov process (in the present example  $X_t = (P_t, G_t)$ ), and by  $\psi_m(t, X_t)$  the respective (time-dependent) payoff rates.

Changing an output level is costly, requiring extra fuel and various overhead costs. We denote by  $C_{i,j}$  the switching costs from state  $i$  to state  $j$ , with potential dependence on time and current state  $C_{i,j}(t, X_t)$ . Dependence on  $X_t$  is very important in practice, where switching typically requires extra use of input fuel so that in our power plant example  $C_{i,j}$  is a function of  $G_t$ . More generally,  $C_{i,j}$  may represent all sorts of operational constraints associated with a dispatch decision. For instance,  $C_{i,j}$  can be used to take into account the time delay effect of having to gradually ‘ramp-up’ and ‘ramp-down’ the turbines, an issue that we revisit more carefully in Section 4.2.2 below.

The switching costs are discrete with  $C_{i,j} > \epsilon > 0$ , for all  $i \neq j$  and some  $\epsilon > 0$ . We take  $C_{i,i} = 0$ , and assume that  $C$  satisfies the triangle inequality

$$C_{i,j} \leq C_{i,k} + C_{k,j}, \quad (2)$$

for any regimes  $i, j$  and  $k$ . This assumption is without loss of generality; one can always re-define  $C_{i,j} = \min_k(C_{i,k} + C_{k,j})$  achieving (2) and without changing the structure of the problem.

## 2.4. Operating Strategies.

Given a power plant described in the previous section, the renter is interested in its optimal use on the finite time horizon  $[0, T]$ , where  $T$  is on the scale of one year. The flexibility of running the plant comes from the ability to schedule the startup and shutdown orders. We model the latter as a control process  $u = (u_t)$  which describes the plant scheduling as a function of time. The control  $u$  is dynamically chosen and adapted to the information filtration  $\mathcal{F}_t^X \triangleq \sigma(X_s : 0 \leq s \leq t)$ . Because the managerial decisions are discrete, we can represent the strategy  $u$  as

$$u = ((\xi_1, \tau_1), (\xi_2, \tau_2), \dots) \quad (3)$$

where  $\xi_k$  taking values in  $\mathbb{Z}_M \triangleq \{0, \dots, M-1\}$  are the successive modes chosen, and  $0 \leq \tau_{k-1} \leq \tau_k \leq T$  are the ( $\mathcal{F}_t^X$ -stopping) switching times. Due to the assumption made in the previous paragraph about subadditive costs, multiple instantaneous switches with  $\tau_k = \tau_{k+1}$  are suboptimal. From a control-theoretic perspective, one should think of  $u$  as an  $\mathcal{F}^X$ -adapted piecewise-constant process where  $u_t = \sum_{\tau_k < t} \xi_k \mathbb{1}_{[\tau_k, \tau_{k+1})}$  denotes the operating mode at time  $t$ .

The total reward up to fixed final time  $T$  for such control  $u$  and scenario  $\omega \in \Omega$  is then

$$H(x, i, [0, T]; u)(\omega) \triangleq \int_0^T \psi_{u_s}(s, X_s(\omega)) ds - \sum_{\tau_k < T} C_{u_{\tau_k^-}, u_{\tau_k}}, \quad X_0 = x, u_0 = i. \quad (4)$$

The first term on the right hand side of (4) describes the cumulative profit corresponding to  $u$  and the second term counts the associated switching costs at each switching time  $\tau_k$ . The payoff rates  $\psi_i$  can be interpreted as either the actual dollar revenue amounts, or alternatively as utility obtained by the agent from consuming the revenue. The latter interpretation allows us to consider simple time-additive risk-aversion of the form  $\mathbb{E}[\int_0^T e^{-rt} u(c_t) dt]$  where  $u$  is the concave utility function and  $c_t = f(X_t)$  is the income stream from the asset. Some care is needed in this case as this income may be negative.

## 2.5. Control Problem.

Let  $\mathcal{U}(t)$  be the set of all allowed scheduling policies on the interval  $[t, T]$ . We will define  $\mathcal{U}(t)$  rigorously in Section 3.1 below. The operational flexibility problem we investigate may now be stated as finding the value function

$$J(t, x, i) = \sup_{u \in \mathcal{U}(t)} J(t, x, i; u), \quad (5)$$

where  $J(t, x, i; u) = \mathbb{E}[H(x, i, [t, T]; u) | X_t = x, u_t = i]$ . As usual in finance,  $\mathbb{E}$  denotes expectation with respect to a risk-neutral pricing measure  $\mathbb{P}$  (again we are agnostic about how  $\mathbb{P}$  is chosen). In full,

$$J(t, x, i) \triangleq \sup_{u \in \mathcal{U}(t)} \mathbb{E} \left[ \int_t^T \psi_{u_s}(s, X_s) ds - \sum_{\tau_k < T} C_{u_{\tau_k^-}, u_{\tau_k}} \middle| X_t = x, u_t = i \right]. \quad (6)$$

The value function  $J(t, x, i)$  represents the conditional maximum expected value from running the plant on  $[t, T]$  given the initial value of  $X_t = x$  and the initial regime  $u_t = i$ . In economic terms,  $J(t, x, i)$  is 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 most importantly characterize the optimal switching policy  $u^*$ , if it exists, that achieves the supremum in (6). The last item is crucial for practical applications where the agent needs an easily implementable dispatch strategy.

Knowing  $J(t, x, i)$  allows one to assign a basic price to the operational flexibility associated with the plant by computing

$$J(t, x, i) - \max_m \mathbb{E} \left[ \int_t^T \psi_m(s, X_s) ds \mid X_t = x \right], \quad (7)$$

which is the difference between expected profit with the flexibility built-in and the expected profit in the best, but fixed, operating mode  $m^*$ . More refined valuation of managerial flexibility can be given by varying the set of allowed dispatch policies  $\mathcal{U}(t)$  in (6) and studying the difference in the resulting value functions.

The control problem (6) is quite non-trivial. It is a special case of stochastic impulse control, see e.g. Øksendal and Sulem (2005). Because the control  $u$  is finite-valued it is called *optimal switching*. In the general form considered here (observe that we have minimal structure imposed on  $\{X_t\}$ ,  $\psi_i$  and  $C_{i,j}$ ) there are few tools for solving it. Nevertheless, in the next section we will exploit the special structure of (6) and provide a robust numerical algorithm for finding  $J$  and  $u^*$ .

## 2.6. Relation to Existing Literature.

Our model is closely aligned with 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  $\tau_k$  and choice/size of project to start  $\xi_k$ , similar to the structure of (3). A good overview and summary of recent results is given in Guo and Pham (2005). Another related problem first appearing in Jeanblanc and Shiryaev (1995) is optimal payout of dividends 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  $\tau_k$  and dividend amounts  $\xi_k$ , again resembling (3).

Directly related to our setting is the literature on optimal switching. Foremost we should mention the recent work by Hamadène and Jeanblanc (2004) whose 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 classical analytic approach to optimal switching originated with Brekke and Øksendal (1994) who considered a geometric Brownian motion for  $\{X_t\}$  and infinite horizon. A recent paper that provides explicit solutions in this special setting is by Pham and Ly Vath (2005). Another similar work is by Yushkevich (2001) who 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 tolling agreements is due to Brennan and Schwartz (1985). They introduced this 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 2003, and references therein). 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 (Dixit 1989). 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. Clearly this induces risk-aversion and reduces NPV. At the same time, because the manager is able to time her decisions, she can minimize her losses in unfavorable conditions, as well maximize them in good ones. As a result she derives extra benefits that are similar to the time premium for American options and form a crucial ingredient of the project NPV.

### 3. Solution Methodology.

#### 3.1. Setup and Assumptions.

We present our results in some generality in order to emphasize the theoretical tools used. Let  $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t^X), \mathbb{P})$  be a stochastic basis. For our driving process we take a  $d$ -dimensional jump-diffusion  $\{X_t\}$  whose dynamics are given by

$$dX_t = \mu(X_t) dt + \sigma(X_t) \cdot dW_t + J(X_t) (dN_t - \lambda_t(X_t) dt), \quad (8)$$

where  $\{W_t\}$  is a standard Wiener process on  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$  and  $\{N_t\}$  is a Poisson process with intensity  $\lambda(X_t)$ . The filtration  $\mathbb{F}$  satisfies the *usual* conditions, with  $\mathcal{F}_0$  being trivial. Each of  $\mu$ ,  $\sigma$ ,  $J$  and  $\lambda$  can be time dependent; we suppress dependence on  $t$  only for convenience. We will write  $\{X_s^{t,x}\}$  to indicate the process conditional on  $X_t = x$ , with law  $\mathbb{P}^{t,x}$ .

Throughout we assume that the stochastic differential equation (SDE) (8) is non-degenerate in the sense that it has a unique strong solution and the (symmetric) diffusion matrix  $a(x) = \sigma^t(x)\sigma(x)$  is positive-definite,  $x \cdot a \cdot x^T > \epsilon > 0$ ,  $\forall x \in \mathbb{R}^d$ . We shall also use the notation

$$\mathcal{S}_T^p \triangleq \{Z; Z_t \in \mathcal{F}_t, \mathbb{E}[\sup_{t \in [0, T]} |Z_t|^p] < \infty\}, \quad p \geq 1$$

and make the following standing assumption.

**ASSUMPTION 1.** *For each  $m = 0, 1, \dots, M - 1$ , the reward function  $\psi_m: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$  is locally Lipschitz in  $(t, x)$  and  $\psi_m(\cdot, X) \in \mathcal{S}_T^2$  uniformly in  $X_0 = x$  restricted to bounded sets.*

The last condition is obviously satisfied if each component of  $X$  is sup-square integrable  $X^n \in \mathcal{S}_T^2$  and all the rewards are of linear growth uniformly in time,  $|\psi_m(t, x)| < C(1 + \|x\|)$ , where  $\|x\| = \sum_i x_i^2$  denotes the Euclidean norm.

We denote by  $\mathcal{U}$  (respectively  $\mathcal{U}(t)$ ) the set of all *acceptable* controls on  $[0, T]$  (resp.  $[t, T]$ ). The basic  $\mathcal{U}$  consists of all adapted, right-continuous with left-limits (rcll),  $\mathbb{Z}_M$ -valued processes  $u$  of a.s. finite variation on  $[0, T]$ . This last requirement implies that the number of switches is finite almost surely. Thus, there exists a random integer  $N_T = N_T(u)$  such that  $\tau_{N_T} = T$ , and  $\mathbb{P}[\tau_k < T \forall k \geq 0] = 0$ . This restriction is superfluous if the expected maximum gain is finite a.s.,  $\mathbb{P}^x[\int_0^T \max_i \psi_i(s, X_s^x) ds = +\infty] = 0$  as infinite number of switches implies infinite switching costs. Further possible constraints for  $\mathcal{U}$  will be discussed later on.

**3.1.1. Canonical Example.** Our canonical example for  $\{X_t\}$  is a  $d$ -dimensional exponential Ornstein-Uhlenbeck process with jumps, namely

$$\begin{aligned} dX_t^n &= X_t^n [\kappa^n (\theta^n - \log X_t^n) dt + \Sigma^n \cdot dW_t + \xi^n dN_t], & n = 1, \dots, d, & (9) \\ \text{or} \quad d(\log X_t^n) &= \kappa^n (\theta^n - \frac{\sigma_n^2}{2\kappa^n} - \log X_t^n) dt + \Sigma^n \cdot dW_t + \xi^n dN_t, & X_0^n = x^n, & \end{aligned}$$

where  $\{W_t\}$  is a  $d$ -dimensional standard Wiener process,  $\sigma_n^2 = (\Sigma^{n,1})^2 + \dots + (\Sigma^{n,d})^2$ ,  $\Sigma = [\Sigma^n]_{n=1, \dots, d} = [\Sigma^{n,p}]_{n,p=1, \dots, d} \in \mathbb{R}^{d \times d}$  is a constant non-degenerate volatility matrix and  $\xi^n \sim \exp(\lambda_n)$ . The last statement means that the  $n$ -th component  $X^n$  experiences jumps of mean size  $\xi_n$ -percent

that occur at constant intensity  $\lambda_n$ . Notice that the dependence between the components of  $X_t$  is only through the matrix  $\Sigma$  (dependence between the jump times can also be added). The model is meant to reflect the widely-documented price spikes and mean-reversion observed in commodity markets. The attractiveness of (9) is that without the jump component,  $\log X_t$  is Gaussian. Because jumps arrive at constant intensity, this means that  $X_t$  can be easily simulated and allows for some explicit computations and approximations. Moreover, in this case Assumption 1 is automatically satisfied for any reward of polynomial growth.

There is still an ongoing debate about appropriate assumptions about electricity prices given the evolving market conditions (including recent sustained price increases). We focus on a mean-reverting  $\{X_t\}$  because commodity prices are expected to be long-term stationary based on a supply-demand equilibrium. More to the point, if  $\{X_t\}$  is not stationary, repeated operational switching is unlikely as the drift would eventually make one regime preferable to all the rest. Thus, mean-reversion is desirable to keep our model ‘interesting’. Empirical studies (Duan and Pliska 2004, Eydeland and Wolyniec 2003) have suggested that the spark spread  $P_t - \overline{HR} \cdot G_t$  is indeed stationary. Another possible alternative (see for instance Carmona and Durrleman (2003)) is to discard  $(P_t)$  and  $(G_t)$  and directly work with a one-dimensional  $\{X_t\}$  that models the spark-spread. This simplifies much of the analysis, but we stress that a single factor is insufficient to fully capture all the effects observed in real markets.

### 3.2. Iterative Optimal Stopping Formulation.

To solve (6) of Section 2.4, we first consider a restricted situation where we put a fixed upper bound on the total number of switches allowed. We define

$$\mathcal{U}^k(t) \triangleq \{u \in \mathcal{U}(t) : \tau_\ell = T \text{ for } \ell \geq k+1\}$$

to be the set of all admissible strategies on  $[t, T]$  with at most  $k$  switches, and we denote by  $\overline{J}^k$  the value function where we optimize only over  $\mathcal{U}^k$ ,

$$\overline{J}^k(t, x, i) = \sup_{u \in \mathcal{U}^k(t)} \mathbb{E} \left[ \int_t^T \psi_{u_s}(s, X_s) ds - \sum_{\ell \geq 1, t \leq \tau_\ell < T} C_{u_{\tau_{\ell-1}}, u_{\tau_\ell}} \middle| X_t = x, u_t = i \right]. \quad (10)$$

We now observe that optimization over  $\mathcal{U}^{k+1}$  and over  $\mathcal{U}^k$  should be related. In fact, Bellman optimality principle says that the problem with  $k+1$  switches is equivalent to finding the first optimal switching time  $\tau$ , which maximizes the initial payoff until  $\tau$ , plus the value function at  $\tau$  corresponding to optimal switching with  $k$  switches. This formal argument suggests that we should be able to solve our control problem through a recursive heap of simpler optimal stopping problems.

For a given stopping time  $\nu$ , let  $\mathcal{S}_\nu \triangleq \{\tau \leq T : \mathbb{F}\text{-stopping time such that } \nu \leq \tau \text{ a.s.}\}$  be the set of all stopping times after  $\nu$ . Thus,  $\mathcal{S} \equiv \mathcal{S}_0$  is the set of all  $\mathbb{F}$ -stopping times bounded by  $T$ . We define  $J^k(t, x, i)$  for  $k = 0, 1, 2, \dots, 0 \leq t \leq T, i \in \mathbb{Z}_M$ , by

$$\begin{aligned} J^0(t, x, i) &\triangleq \mathbb{E} \left[ \int_t^T \psi_i(s, X_s) ds \middle| X_t = x \right], \\ J^k(t, x, i) &\triangleq \operatorname{ess\,sup}_{\tau \in \mathcal{S}_t} \mathbb{E} \left[ \int_t^\tau \psi_i(s, X_s) ds + \mathcal{M}^{k,i}(\tau, X_\tau) \middle| X_t = x \right], \quad k \geq 1. \end{aligned} \quad (11)$$

We shall see below that the essential supremum in (11) is in fact a supremum, which is moreover attained, so it is actually a maximum. The intervention operator  $\mathcal{M}$  reflects the recursive nature



of (11) and specifies the best value that can be achieved by making an immediate switch from current mode  $i$  given  $k$  switches remaining,

$$\mathcal{M}^{k,i}(t, x) \triangleq \max_{j \neq i} \left\{ -C_{i,j} + J^{k-1}(t, x, j) \right\}. \quad (12)$$

If there are only two regimes like in Hamadène and Jeanblanc (2004), the maximum in (12) is not needed as there is only one regime to switch into.

The rest of the section is devoted to proving regularity properties of  $J^k(t, x, i)$  and showing that the construction (11) provides the means of tackling the original problem (6). Observe that (11) defines  $J^k$  in terms of an optimal stopping problem, a much simpler setting than the stochastic control formulation of (6).

We first check that (11) is well-posed and in particular that  $J^k$  can be chosen to be jointly measurable. To do so we recall the theory of Snell envelopes from general optimal stopping (El-Karoui 1981). The results we need are summarized in the following proposition.

**PROPOSITION 1.** (*Hamadène and Lepeltier 2000*) *Let  $Z$  be an  $\mathbb{R}$ -valued rcll, not necessarily Markov, process adapted to the filtration  $(\mathcal{F}_t)$ . We assume that  $Z$  is left-continuous in expectation and has square-integrable supremum  $Z \in \mathcal{S}_T^2$ . For each  $\nu \in \mathcal{S}$ , define the nonnegative random variable*

$$Y(\nu) \triangleq \operatorname{ess\,sup}_{\tau \in \mathcal{S}_\nu} \mathbb{E} \left[ Z_\tau \mid \mathcal{F}_\nu \right]. \quad (13)$$

Then,

1. *The family  $\{Y(\nu)\}$  admits a version which is a rcll  $(\mathcal{F}_t)$ -supermartingale  $Y^r$  satisfying  $Y(\nu) = Y_\nu^r$ , namely  $Y^r$  stopped at time  $\nu$ .  $Y^r$  is called the Snell envelope of  $Z$ .*

2. *The minimal optimal stopping time  $\tau^*$  for  $Y(0)$  exists and is explicitly given by*

$$\tau^* = \inf \left\{ s \geq 0 : Y_s^r \leq Z_s \right\}.$$

Thus,  $Y_0^r = \mathbb{E}[Z_{\tau^*}] = \sup_{\tau} \mathbb{E}[Z_\tau]$ .

3.  *$Y^r$  is a supermartingale on  $[0, T]$  and a martingale on the random interval  $[0, \tau^*)$ . Moreover,  $Y^r \in \mathcal{S}_T^2$ .*

4. *If  $Z$  is upper semicontinuous from the left and of class [D] (that is  $\{Z_\tau : \tau \in \mathcal{S}\}$  is uniformly integrable) then  $Y^r$  is of class [D].*

Going back to our setting, pick an initial value  $X_0 = x$  and 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 Proposition 1. Then  $Y^{k,i}$ , the Snell envelope of  $Z^{k,i}$ , solves  $Y_t^{k,i} = \operatorname{ess\,sup}_{\tau \in \mathcal{S}_t} \mathbb{E}^x [Z_\tau^{k,i} \mid \mathcal{F}_t]$  and comparing with (11) we see that

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

Accordingly,  $J^k(t, X_t^x, i)$  is a supermartingale, is square-integrable, and, if  $X$  satisfies item 4 of Proposition 1, is continuous. This identification also shows that the minimal optimal stopping time  $\tau_k^*$  corresponding to the Snell envelope defined by  $J^k(t, x, i)$  is simply

$$\tau_k^* = \inf \{ s \geq t : J^k(s, X_s^{t,x}, i) = \mathcal{M}^{k,i}(s, X_s^{t,x}) \} \wedge T. \quad (14)$$

Armed with these facts we next show that the recursive definition is equivalent to the original problem.

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

*Proof:* The theorem states that a ‘global’ optimization using a fixed maximum of  $k$  switches in (10) is equivalent to a successive ‘local’ optimization using one switch at a time in (11). The proof is based on direct use of the properties of Snell envelope and proceeds by induction on the number of switches left. The case  $k = 0$  is immediate. Indeed,  $\overline{J}^0 = J^0$ , since the sum appearing in the right hand side of definition (10) is not present for  $k = 0$ . Take  $k \geq 1$  and assume that  $\overline{J}^{k-1} = J^{k-1}$ . Let  $u = ((\xi_1, \tau_1), (\xi_2, \tau_2), \dots, (\xi_k, \tau_k))$  be any given strategy in  $\mathcal{U}^k(t)$ . Then

$$\begin{aligned}
J^k(t, x, i) &\geq \mathbb{E} \left[ \int_t^{\tau_1} \psi_i(s, X_s^{t,x}) ds + \mathcal{M}^{k,i}(\tau_1, X_{\tau_1}^{t,x}) \Big| X_t = x \right] & (15) \\
&\geq \mathbb{E} \left[ \int_t^{\tau_1} \psi_i(s, X_s^{t,x}) ds - C_{i, \xi_1} + J^{k-1}(\tau_1, X_{\tau_1}^{t,x}, \xi_1) \Big| X_t = x \right] \\
&\geq \mathbb{E} \left[ \int_t^{\tau_1} \psi_i(s, X_s^{t,x}) ds - C_{i, \xi_1} + \overline{J}^{k-1}(\tau_1, X_{\tau_1}^{t,x}, \xi_1) \Big| X_t = x \right] \\
&\geq \mathbb{E} \left[ \int_t^{\tau_2} \psi_{u_s}(s, X_s^{t,x}) ds - C_{i, \xi_1} - C_{\xi_1, \xi_2} + J^{k-2}(\tau_2, X_{\tau_2}^{t,x}, \xi_2) \Big| X_t = x \right] \\
&\vdots \\
&\geq \mathbb{E} \left[ \int_t^{\tau_k} \psi_{u_s}(s, X_s^{t,x}) ds - \sum_{1 \leq \ell \leq k} C_{\xi_{\ell-1}, \xi_\ell} + \int_{\tau_k}^T \psi_{\xi_k}(s, X_s^{t,x}) ds \Big| X_t = x \right] \\
&= \overline{J}^k(t, x, i; u), & (16)
\end{aligned}$$

with the convention  $\xi_0 = i$ . To go from (15) to (16) we successively used the defining property (11) of  $J^k$ , the definition of the intervention operator  $\mathcal{M}^{k,i}$ , the fact that the induction hypothesis implies  $J^{k-1}(\tau_\ell, X_{\tau_\ell}, \xi_\ell) = \overline{J}^{k-1}(\tau_\ell, X_{\tau_\ell}, \xi_\ell)$ , and the definition (10) of  $\overline{J}^k$ . To prove that we have equality between  $J^k$  and  $\overline{J}^k$ , we notice that all the inequalities in (15) become equalities if we use the control  $u^* = ((\xi_1^*, \tau_1^*), \dots, (\xi_k^*, \tau_k^*)) \in \mathcal{U}^k(t)$  defined by  $\tau_0^* = 0$ , the stopping times

$$\tau_\ell^* \triangleq \inf \left\{ s \geq \tau_{\ell-1}^* : J^\ell(s, X_s, i) = \max_{j \neq i} (-C_{i,j} + J^{\ell-1}(s, X_s, j)) \right\} \wedge T, \quad (17)$$

for  $\ell = 1, \dots, k$ , and the corresponding sequence of regimes  $\xi_\ell^* \triangleq \arg \max_j \mathcal{M}^{\ell,i}(\tau_\ell^*, X_{\tau_\ell^*})$ .  $\square$

Observe that as an aside, the proof of Theorem 1 furnishes the *existence* of an optimal switching policy for  $\overline{J}^k$  and directly demonstrates that this policy can be chosen to be of the ‘barrier’ type, i.e. in terms of hitting times of  $\{X_t\}$ .

On an economic level, the value functions  $J^k$  provide a fine-level granularity with respect to the operational flexibility. In practice, some assets have stringent cycling restrictions as frequent switches increase maintenance cost and potential tear-and-wear. Hence, before signing the tolling agreement, the agent can use the marginal differences  $J^{k+1} - J^k$  to precisely value each degree of contractual flexibility.

Our next theorem shows that as expected, letting  $k$  increase without bound, the restricted problem (10) converges to the original unrestricted setting of (6). This idea of representing impulse control problems as limits of sequences of optimal multiple stopping problems goes back to 1980s, see e.g. Lepeltier and Marchal (1984), Øksendal and Sulem (2005) and more recently Hamadène and Jeanblanc (2004). What is new, however, is its use for numerical computations as we do in Section 4.

**THEOREM 2.**  $\lim_{k \rightarrow \infty} J^k = J$  *pointwise*.

*Proof:* Recall that  $J(t, x, i)$  was defined in (6) and  $J^k(t, x, i)$  in (11). Since having more switches is always advantageous,  $J^{k+1} \geq J^k$  and the limit, say  $J^\infty$ , exists pointwise. Clearly,  $J^\infty \leq J$  since  $J \geq \overline{J}^k = J^k$  for every integer  $k \geq 1$ . It remains to show that  $J^\infty \geq J$ . Let  $\epsilon > 0$  be arbitrary,

and let us assume that  $u^* = ((\xi_1^*, \tau_1^*), \dots)$  is an  $\epsilon$ -optimal policy for  $J(t, x, i)$  in the sense that  $J(t, x, i; u^*) \geq J(t, x, i) - \epsilon$ . For any integer  $k \geq 1$  we then define a policy  $u^{k,*} \in \mathcal{U}^k(t)$  such that  $u_t^{k,*} = u_t^*$  if  $t \leq \tau_k^*$ , and  $u_t^{k,*} = u_{\tau_k^*}^*$  otherwise. In other words,  $u^{k,*}$  coincides with  $u^*$  up to the  $k$ -th switch, after which it remains constant. Now

$$\begin{aligned} J(t, x, i; u^*) - J(t, x, i; u^{k,*}) &= \mathbb{E} \left[ \int_{\tau_{k+1}^*}^T (\psi_{u_s^*}(s, X_s) - \psi_{u_{\tau_k^*}^*}(s, X_s)) ds - \sum_{\ell > k} C_{u_{\tau_\ell^*}^*, u_{\tau_\ell^*}^*} \right] \\ &\leq 2 \mathbb{E}[T - \tau_k^*] \cdot \mathbb{E} \left[ \sup_{0 \leq t \leq T} \max_i |\psi_i(t, X_t)| \right] \end{aligned} \quad (18)$$

and the right hand side converges to zero because by admissibility  $\tau_k^* \rightarrow T$  a.s.. It follows that  $\lim_{k \rightarrow \infty} J(t, x, i; u^{k,*}) = J(t, x, i; u^*)$ . However, since  $u^{k,*} \in \mathcal{U}^k(t)$ ,  $J^\infty(t, x, i) \geq J(t, x, i; u^{k,*})$  for any  $k$ . We conclude that

$$J^\infty(t, x, i) \geq \lim_{k \rightarrow \infty} J(t, x, i; u^{k,*}) = J(t, x, i; u^*) \geq J(t, x, i) - \epsilon,$$

which implies  $J^\infty(t, x, i) \geq J(t, x, i)$  since  $\epsilon$  was arbitrary.  $\square$

Thanks to our recursive construction, it also becomes clear that because  $\{X_t\}$  is Markov, the optimal policy of  $J^k$  can be chosen to be Markovian as well (Lepeltier and Marchal 1984). Indeed, each switch by itself is Markovian in (17). From the proof of Theorem 2 it follows that for any  $\epsilon > 0$  we can find an  $\epsilon$ -optimal Markovian strategy for  $J(t, x, i)$ . Of course the ability to make optimal dispatching decisions based solely on today's information is intuitive, but a rigorous analytic proof is often hard to come by. Moreover, the Markovianity of optimal strategies implies that the formal Dynamic Programming Principle holds for the value function, i.e. for any  $t' \geq t$ ,

$$J(t, x, i) = \sup_{u \in \mathcal{U}(t, t')} \mathbb{E} [H(x, i, [t, t']; u) + J(t', X_{t'}^{t,x}, u_{t'}) | X_t = x, u_t = i]. \quad (19)$$

As a final result we now directly show that for reasonable models the value function is continuous in  $x$ .

LEMMA 1. *Suppose the flow of  $(X_t)$  is Lipschitz, namely that  $\mu$  and  $\sigma$  in (8) are continuously differentiable with bounded derivatives. Then the value functions  $J(t, x, i)$  are Lipschitz continuous in  $x$ .*

*Proof:* By a standard result, the assumptions of the lemma imply that for any  $p > 1$ , there is a  $C_X$  such that  $\mathbb{E}^x[\sup_t \|X_t\|^p] \leq C_X(1 + \|x\|^p)$  (Protter 2004, p. 342). The rest follows by a simple estimate and Gronwall's inequality:

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

$\square$

### 3.3. Continuity of the Switching Boundary.

LEMMA 2. *Suppose that Assumption 1 holds and that the payoff rates  $\psi_i$  and the process  $\{X_t\}$  are time-homogeneous. Then the value function  $J(t, x, i)$  is Lipschitz continuous in  $t$ , locally uniformly in  $x$ .*

*Proof:* To emphasize the time-interval considered, let us write  $J(t, x, i; T)$  for the value function of (6). By time-homogeneity,  $J(t, x, i; T) = J(0, x, i; T - t)$ , i.e. we can shift the problem to calendar time zero and maturity  $(T - t)$ . Therefore for  $t < t'$ ,

$$\begin{aligned} |J(t, x, i) - J(t', x, i)| &= |J(0, x, i; T - t) - J(0, x, i; T - t')| \\ &\leq \sup_{u \in \mathcal{U}(T-t')} \mathbb{E}^x \left[ \int_{T-t'}^{T-t} |\psi_{u_s}(X_s^x)| ds \right] \\ &\leq (t' - t) \mathbb{E} \left[ \sup_{0 \leq s \leq T} \max_i |\psi_i(X_s^x)| \right] \end{aligned} \quad (20)$$

and the right hand side is uniform in  $x$  restricted to a bounded set by Assumption 1.  $\square$

At this stage it is important to remark that, if it is optimal to switch at a given time  $\tau$  from mode  $i$  to mode  $j$ , then for all the modes  $k$  such that

$$C_{i,j} = C_{i,k} + C_{k,j} \quad (21)$$

it is also optimal to switch, at the same time  $\tau$ , to site  $j$ . Indeed, optimality of the switch at time  $\tau$  from mode  $i$  to mode  $j$  implies that

$$-C_{i,j} + J(\tau, X_\tau, j) > -C_{i,k} + J(\tau, X_\tau, k)$$

and because of the additivity (21), we obtain  $-C_{k,j} + J(\tau, X_\tau, j) > J(\tau, X_\tau, k)$ , which shows that switching to mode  $j$  is better than remaining in mode  $k$  at time  $\tau$ .

The above argument implies that when the rewards are monotone in each individual component  $X^n$  of the  $(X_t)$ -vector and the switching costs are additive then under mild regularity assumptions on  $J(t, x, i)$  the switching regions will be connected. Indeed, suppose we can find an ordering of the regimes such that for any  $i > j$ ,

$$\frac{\partial(\psi_i - \psi_j)}{\partial x^n} \geq 0 \quad (\text{respectively } \leq 0). \quad (22)$$

Then it can be shown that  $\partial_{x^n}(J(t, x, i) - J(t, x, j)) \geq 0$  (resp.  $\leq 0$ ), since the switching times of  $J(t, x, j)$  are also switching times of  $J(t, x, i)$  due to the cost additivity. Hence, the difference between the two value functions with different regimes is monotone in  $x^n$  and the set

$$\text{Switch}(i, j; t) \triangleq \left\{ x \in \mathbb{R}^d : -C_{i,j} + J(t, x, j) > \max_{j' \neq j} \{-C_{i,j'} + J(t, x, j')\} \right\} \quad (23)$$

of all  $x$ 's where it is optimal to immediately switch from regime  $i$  to  $j$  is connected in this case. The switching boundaries are the boundaries of  $\text{Switch}(i, j; t)$  as a function of  $t$  for different pairs  $(i, j)$ . By Lemma 2, the value function  $J$  is continuous in  $t$ , and consequently so is the switching set and its boundary. Condition (22) on the payoff rates is quite reasonable in practice, for instance all three of our examples in Section 5 satisfy it.

### 3.4. Comparison to the Quasi-Variational Formulation.

The classical method of solving impulse control problems driven by Markov processes goes back to the fundamental work of Bensoussan and Lions (1984). The idea is that the state space is the union of the continuation region where  $\{X_t\}$  is left uncontrolled, and the exercise regions where the operational flexibility is immediately exercised. Let  $\mathcal{L}_X$  be the infinitesimal generator of  $\{X_t\}$  and let us denote by  $\mathcal{M}\phi(t, x, i) = \max_{j \neq i} \{-C_{i,j} + \phi(t, x, j)\}$  the intervention operator. Then roughly speaking (see Brekke and Øksendal (1994) for full details), one needs to study the *quasi-variational inequality*

$$\max \left( \mathcal{L}_X \phi(t, x, i) + \psi_i(t, x), -\phi(t, x, i) + \mathcal{M}\phi(t, x, i) \right) = 0. \quad (24)$$

In other words, the function  $\phi$  must dominate the barrier and satisfy the Kolmogorov PDE inside the continuation region. The so-called verification theorem states that a sufficiently smooth solution  $\phi(t, x, i)$  of (24) is equal to the value function  $J(t, x, i)$  of (6). More generally, it can be shown that the value function is the unique *viscosity* solution of (24) (Pham 2005). However, working with viscosity solutions is laborious; instead the ‘smooth pasting’ condition is usually invoked asserting that  $J(t, x, i)$  is differentiable across the respective free boundaries. See for example Pham (2005) and the references therein for a discussion of smooth pasting. In the literature this assumption is often made by invoking heuristic arguments and even then it only leads to a system of implicit equations with the existence-uniqueness of a solution (i.e. a ‘correct’ continuation region) still not directly verifiable. At the risk of being unfair, we could say that the variational method is a theoretical overkill that poses too many technical challenges in practice. In contrast our probabilistic formulation sidesteps all these difficulties; in particular we do not need to make any a priori assumptions about the smoothness of  $J(t, x, i)$ .

Secondly, the analytic formulation above is poorly suited for numerical implementation. In particular, computational difficulties are encountered when  $X_t$  is multi-dimensional and/or when  $X_t$  contains jumps, in which case  $\mathcal{L}_X$  becomes an integro-differential operator.

### 3.5. Reflected Backward SDEs.

The recently developed theory of reflected backward stochastic differential equations (BSDE’s) gives another useful characterization of the solution to (6). Fix an initial condition  $X_0 = x$  and let  $Y_t^{k,i} \triangleq J^k(t, X_t^x, i)$ . Note that  $Y_t^{k,i}$  is an  $\mathcal{F}_t^X$ -measurable random quantity. Then  $Y_t^{k,i}$  solves

$$\begin{cases} Y_t^{k,i} = \int_t^T \psi_i(s, X_s^x) ds + A_T^{k,i} - A_t^{k,i} - \int_t^T Z_s^{k,i} \cdot dW_s, \\ Y_t^{k,i} \geq \mathcal{M}^{k,i}(t, X_t^x), \quad \text{as defined in (12),} \\ \int_0^T (Y_t^{k,i} - \mathcal{M}^{k,i}(t, X_t^x)) dA_t^{k,i} = 0, \quad A_0^{k,i} = 0, \end{cases} \quad (25)$$

The auxiliary process  $Z^{k,i}$  satisfies  $\mathbb{E}[\int_0^T \|Z_t^{k,i}\|^2 dt] < \infty$  and can be interpreted as a kind of conditional expectation. The second auxiliary process  $A^{k,i}$  is a compensator (continuous, increasing and square-integrable) that increases only when  $Y^{k,i}$  hits the barrier  $\mathcal{M}^{k,i}$ . Let us also mention that the integrability assumption  $\mathcal{M}^{k,i} \in \mathcal{S}_T^2$  that we checked earlier is necessary for (25) to make sense.

The elegance of (25) can be seen from the result of (El-Karoui et al. 1997, Prop. 2.3), who assert that subject to the above integrability conditions, there is a unique triple  $(Y^{k,i}, Z^{k,i}, A^{k,i})$  solving (25) and necessarily satisfying  $Y_0^{k,i} = J^k(0, x, i)$ . In fact, denoting  $Y_s^{k,i,t,x}$  the solution of (25) started with  $X_t = x$ , there exists a deterministic and jointly continuous function  $u^{k,i}(t, x)$  such that  $u^{k,i}(t, x) = Y_t^{k,i,t,x}$ . This  $u^{k,i}$  is known (El-Karoui et al. 1997) to be the unique viscosity solution of the QVI (24).

The BSDE formulation clarifies the continuity and stability properties of the value function. Indeed, if the barrier  $\mathcal{M}^{k,i}$  is continuous in  $t$ , it is clear that the solution  $Y^{k,i}$  of (25) is also continuous. Furthermore, analogues of stochastic flow theory imply that under regularity conditions on the forward SDE (8) giving the dynamics of  $\{X_t\}$ ,  $Y^{k,i}$  will be continuously differentiable with respect to the initial condition  $x$ , and other parameters of (8). By induction the same is true of the recursive value functions  $J^k(t, x, i)$  of our problem, so that the Greek sensitivities (with respect to asset prices, underlying volatilities, etc.) of the operational flexibility problem are well-defined.

REMARK 1. The original value function  $J(t, x, i)$  can be formally represented as a coupled system of reflected BSDE's for  $(Y^i)_{i \in \mathbb{Z}_M}$ ,  $Y_t^i \equiv J(t, X_t^x, i)$ :

$$\begin{cases} Y_t^i = \int_t^T \psi_i(s, X_s) ds + A_T^i - A_t^i - \int_t^T Z_s^i \cdot dW_s, \\ Y_t^i \geq \max_{j \neq i} \{-C_{i,j} + Y_t^j\}, \quad \int_0^T (Y_t^i - \max_{j \neq i} \{-C_{i,j} + Y_t^j\}) dA_t^i = 0, . \end{cases} \quad (26)$$

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 (2004) explicitly proved existence of solutions to (26) by working with the difference process  $Y^1 - Y^2$ . For  $M > 2$  the problem remains open.

## 4. Numerical Solution.

Having reduced the optimal scheduling problem to a cascade of recursive optimal stopping problems appearing in (11), we now show that this transformation also leads to numerical methods that give an efficient and robust way of computing  $J(t, x, i)$ .

### 4.1. Dynamic Programming in Discrete Time.

In order to do numerical computation, we first pass from the continuous time to discrete time. Let  $\mathcal{S}^\Delta = \{m\Delta t, m = 0, 1, \dots, M^\sharp\}$  be a discrete time grid with  $\Delta t = T/M^\sharp$ . Switches are now allowed only at grid points, i.e.  $\tau_k \in \mathcal{S}^\Delta$  and the new restricted set of admissible strategies is labelled  $\mathcal{U}^\Delta$ . This limiting of managerial flexibility is similar to looking at Bermudan options as approximations to American exercise rights.

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 of Bellman. 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, (11) becomes

$$\begin{aligned} J^k(t_1, X_{t_1}, i) &= \max \left\{ \mathbb{E} \left[ \int_{t_1}^{t_2} \psi_i(s, X_s) ds + J^k(t_2, X_{t_2}, i) \mid \mathcal{F}_{t_1} \right], \mathcal{M}^{k,i}(t_1, X_{t_1}) \right\} \\ &\simeq \left( \psi_i(t_1, X_{t_1}) \Delta t + \mathbb{E} [J^k(t_2, X_{t_2}, i) \mid \mathcal{F}_{t_1}] \right) \vee \left( \max_{j \neq i} \{-C_{i,j} + J^{k-1}(t_1, X_{t_1}, j)\} \right). \end{aligned} \quad (27)$$

We see that to solve the problem it suffices to have a computationally efficient algorithm for evaluating the *conditional expectations* appearing in (27).

Recall that the conditional expectation  $\mathbb{E}[f(X_{t_2}) \mid \mathcal{F}_{t_1}]$  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  $x \mapsto E_{t_1}(x) \triangleq \mathbb{E}[f(X_{t_2}) \mid X_{t_1} = x]$  and our strategy is to approximate this map. This is a well-known statistical problem; here we choose to concentrate on an approach first described by Longstaff and Schwartz (2001) and Tsitsiklis and Van Roy (2001). Many alternatives are available and we review and compare them in Section 5.2.

### 4.2. Regression Monte-Carlo Method.

The idea of the regression Monte Carlo approach is to *project*  $E_t$  onto a truncation of a basis of the Hilbert space  $L^2(\mathbb{P})$ :

$$E_t(x) \simeq \hat{E}_t(x) = \sum_{j=1}^{N_B} \alpha_j B_j(x), \quad (28)$$

where  $B_j(x)$  are the  $N_B$  bases and  $\alpha_j$  the  $\mathbb{R}$ -valued coefficients. The projection can now be approximated with an empirical regression. The algorithm generates a large Monte Carlo sample  $(x_{t_1}^n, x_{t_2}^n)_{n=1}^{N_p}$  from the joint distribution of  $X_{t_1}$  and  $X_{t_2}$ . The empirical values  $f(x_{t_2}^n)$  are then regressed against  $\{B_j(x_{t_1}^n)\}$  to obtain the (sample) coefficients  $\alpha_j$  after which we can use (28) to compute the conditional expectations.

Our contribution is to extend this approach to simultaneously compute all the  $J^k$ 's. Note that this is not a trivial task since we face a heap of inter-dependent and recursive optimal stopping problems. Begin by generating  $N_p$  sample paths  $x^n$  of the discretized  $\{X_t\}$  process with a fixed initial condition  $X_0 = x = x_0^n$ . We will approximate  $J^k(0, x, i)$  by  $\frac{1}{N_p} \sum_n J^k(0, x_0^n, i)$ . The pathwise values  $J^k(t, x_t^n, i)$  are computed recursively in a backward fashion, starting with  $J^k(T, x_T^n, i) = 0$ . To implement (27), for a given step  $t_1 = m\Delta t$  and regime  $i$  we regress the known future values  $J^k((m+1)\Delta t, x_{(m+1)\Delta t}^n, i)$  onto the current underlying values  $\{B_j(x_{m\Delta t}^n)\}$ . As a result of regression we obtain a prediction  $\hat{E}_{m\Delta t}[J^k(m\Delta t, \cdot, i)](x_{m\Delta t}^n)$  for the continuation value along the  $n$ -th path. Comparing this with the current value  $J^{k-1}(m\Delta t, x_{m\Delta t}^n, j)$  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  $J^{k-1}(m\Delta t, \cdot, \cdot)$  is known when computing  $J^k(m\Delta t, \cdot, \cdot)$ . The efficiency is maintained because we use *the same set* of paths to compute all the 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 accumulate with respect to number of switches.

It turns out that for numerical efficiency rather than directly computing the value function, it is much better to instead keep track of the optimal stopping times. This was the beautiful insight of Longstaff and Schwartz (2001). 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

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

Then we have the analogue of (27) for  $\tau^k$ :

$$\tau^k(m\Delta t, x_{m\Delta t}^n, i) = \begin{cases} \tau^k((m+1)\Delta t, x_{(m+1)\Delta t}^n, i), & \text{no switch;} \\ m, & \text{switch,} \end{cases} \quad (29)$$

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

$$\hat{j}^n(t_1; i) = \arg \max_j \left( -C_{i,j} + J^{k-1}(t_1, x_{t_1}^n, j), \psi_i(t_1, x_{t_1}^n) \Delta t + \hat{E}_{t_1} [J^k(t_2, \cdot, i)](x_{t_1}^n) \right). \quad (30)$$

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

$$J^k(m\Delta t, x_{m\Delta t}^n, i) = \begin{cases} \psi_i(m\Delta t, x_{m\Delta t}^n) \Delta t + J^k((m+1)\Delta t, x_{(m+1)\Delta t}^n, i), & \text{no switch;} \\ -C_{i,j} + J^{k-1}(m\Delta t, x_{m\Delta t}^n, j), & \text{switch to } j. \end{cases} \quad (31)$$

Observe that in this version the regression is used solely to update the optimal stopping times  $\tau^k$  and the regressed values are never stored directly. This helps to eliminate potential biases from the regression step. In particular, as long as correct switching decisions are made, the regression approximation error never enters  $J^k(t, X_t, i)$ .

The above description of the algorithm made direct use of the restricted value functions  $J^k$ . This is necessary if there is an upper operational bound on the number of total switches. If there is not, we can simplify by computing  $J(t, x, i)$  directly. Namely, we simply replace all  $J^k$ 's in (27)-(31)

with just  $J$ . Hence, we regress  $J((m+1)\Delta t, x_{(m+1)\Delta t}^n, i)$  against  $\{B_j(x_{m\Delta t}^n)\}$  and use the resulting map  $\hat{E}_{t_1}$  to replace (30) with

$$\hat{J}^n(t_1; i) = \arg \max_j \left( -C_{i,j} + J(t_1, x_{t_1}^n, j), \psi_i(t_1, x_{t_1}^n) \Delta t + \hat{E}_{t_1} [J(t_2, \cdot, i)](x_{t_1}^n) \right).$$

We stress that this is not equivalent to computing  $J^K(t, x, i)$  for large  $K$  in (31) because the values used in the regression are different. Numerically, computing  $J(t, x, i)$  directly is definitely much faster and also appears to be more stable, however (31) is conceptually clearer and more tractable analytically.

**4.2.1. 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 in the original paper of Longstaff and Schwartz (2001), the indicator functions of a partition of the domain of  $\{X_t\}$  in Gobet et al. (2005), and the logistic basis in Haugh and Kogan (2004). However, there is now widespread consensus that the numerical precision can be greatly improved by *customizing* the basis (for example, the customization advantage is documented by Andersen and Broadie (2004)). In particular, it helps to use basis functions that resemble the expected shape of the value function. In our examples  $\psi_i$  are usually linear and our favorite set of bases  $B_j(x)$  is of the form  $x^p$ ,  $e^{\alpha x}$  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 suggests that the required number of basis functions is 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.

**4.2.2. 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 *ramping* time. To model this phenomenon we may want to assume that each switch has a delay of  $\delta$ . The control  $u_t$  is still  $\mathcal{F}_t$ -measurable, but the actual regime switch and the corresponding payoff are only realized at  $t + \delta$ . Unfortunately, delay introduces cumbersome technical difficulties, see e.g. Bar-Ilan et al. (2002).

As an approximation of time delay we impose *time separation*, i.e. the constraint  $\tau_k \geq \tau_{k-1} + \delta$ . Consequently, a dispatch decision has immediate effect but then the plant becomes locked-up for a period of  $\delta$ . Time separation is less risky than true delay because decisions have instantaneous impact, however it does limit the flexibility of the operator. A related situation (but for entirely different economic reason) occurs in the case of power Swing options, see Carmona and Touzi (2005). Strictly speaking, with separation the value function is no longer Markov: if a switch is made at  $t$ , time advances to  $t + \delta$  while the agent's hands are tied. Hence, (11) only holds at the switch times. However, this does not pose a practical difficulty since we only compute  $J(m\Delta t, \cdot)$  anyway; the advantage is that in discrete time, especially if  $\delta$  is a multiple of  $\Delta t$ , separation is very convenient to implement. Instead of evaluating  $\mathbb{E}[J^k(t + \Delta t, X_{t+\Delta t}, i) | \mathcal{F}_t]$  we now compute  $\mathbb{E}[J^k(t + \delta, X_{t+\delta}, i) | \mathcal{F}_t]$  the conditional expectation  $\delta/\Delta t$  steps ahead, which is as easy as the original computation.

### 4.3. Summary of Algorithm.

For the reader's convenience we summarize our basic algorithm that combines the Longstaff-Schwartz features together with simultaneous computations of the 'layers'  $J^k$  for  $k = 1, 2, \dots, \bar{K}$ .

1. Select a set of basis functions ( $B_j$ ) and algorithm parameters  $\Delta t, M^\#, N_p, \bar{K}, \delta$ .



2. Generate  $N_p$  paths of the driving process:  $\{x_{m\Delta t}^n, m = 0, 1, \dots, M^\sharp, n = 1, 2, \dots, N_p\}$  with fixed initial condition  $x_0^n = x_0$ .
3. Initialize the value functions and switching times  $J^k(T, x_T^n, i) = 0, \tau^k(T, x_T^n, i) = M^\sharp, \forall i, k$ .
4. Moving backward in time with  $t = m\Delta t, m = M^\sharp, \dots, 0$  repeat the Loop:
  - (a) Compute inductively the layers  $k = 0, 1, \dots, \bar{K}$  using (11). To evaluate the conditional expectation  $\mathbb{E}[J^k(m\Delta t + \Delta t, \cdot, i) | \mathcal{F}_{m\Delta t}]$  regress  $\{J^k(m\Delta t + \Delta t, x_{m\Delta t + \Delta t}^n, i)\}$  against current set of selected basis functions  $\{B_j(x_{m\Delta t}^n)\}_{j=1}^{N_B}$ . Add the reward  $\psi_i(m\Delta t, x_{m\Delta t}^n) \cdot \Delta t$  to the continuation value.
  - (b) Update the switching times and value functions using (29), (30) and (31).
5. end Loop.
6. Check whether  $\bar{K}$  switches are enough by comparing  $J^{\bar{K}}$  and  $J^{\bar{K}-1}$  (they should be equal).

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 4b directly uses the conditional expectations,

$$J^k(m\Delta t, x_{m\Delta t}^n, i) = \mathcal{M}^{k,i}(m\Delta t, x_{m\Delta t}^n) \vee \left[ \hat{E}_{m\Delta t} [J^k((m+1)\Delta t, \cdot, i)](x_{(m+1)\Delta t}^n) + \psi_i(m\Delta t, x_{m\Delta t}^n) \cdot \Delta t \right]. \quad (32)$$

**4.3.1. 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 steps used. Observe that during the main loop we only need to store the buffer  $J(t, \cdot), \dots, J(t + \delta, \cdot)$ ; and  $\tau(t, \cdot), \dots, \tau(t + \delta, \cdot)$ . Consequently, the memory requirements are  $O(N_p \cdot (M^\sharp + D \cdot \bar{K} \cdot M))$  where  $D = \frac{\delta}{\Delta t} + 1 \geq 2$  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  $N_p \times N_B$  where  $N_B$  denotes the dimension of the projection space used for regression purposes. 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 minimal values such as  $M^\sharp = 500, N_p = 25,000$  this is already 50MB of storage.

#### 4.4. Convergence.

The algorithm described in the previous section incorporates several layers of approximations. Four sources of errors can be identified: error due to discretizing the SDE (8), error due to restricting switching times to  $\mathcal{S}^\Delta$ , projection error and Monte Carlo sampling error. To understand the associated effects we shall systematically address each of these and focus on error bounds in terms of  $\Delta t$  and the number of paths  $N_p$ . To do so we return to the framework of backward stochastic differential equations from Section 3.5. The basic TvR scheme (32) 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. Chevance (1997), Bouchard and Touzi (2004), Gobet et al. (2005). Empirical evidence strongly suggests that the 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 (25) and constructs a discrete time approximation to the pairs  $(Y^{k,i}, Z^{k,i}), k = 1, 2, \dots$ . Recall that  $Y_t^{k,i} = J^k(t, X_t^x, i)$ , while  $Z^{k,i}$  will be replaced with a conditional expectation similar to (27).

To simplify the proofs we shall impose,

**ASSUMPTION 2.** *The functions  $x \mapsto \mu(x)$  and  $x \mapsto \sigma(x)$  appearing in (8) are  $\mathcal{C}^2(\mathbb{R}^d)$  with uniformly bounded first derivatives. The payoff functions  $\psi_m(t, x)$  are time-homogeneous, and continuously differentiable with uniformly bounded derivatives.*

The convergence analysis will proceed in several steps, isolating each source of error identified above. Recall that  $M^\#$  and  $N = N_p$  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  $n$ . For brevity,  $C$  is a generic constant that may change from line to line, and  $t_1$  and  $t_2$  are two generic consecutive time grid points  $t_1 = m\Delta t$  and  $t_2 = t_1 + \Delta t$ . Finally, we use the standard notation  $X \vee Y \triangleq \max(X, Y)$  and  $\|X\|_p = \mathbb{E}[|X|^p]^{1/p}$ ,  $p \geq 1$ . At this point let us also recall the following discrete version of Gronwall's inequality:

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

$$a_m \leq e^{C(T-m\Delta t)} \left[ a_{M^\#} + \sum_{j \geq m} b_j + \max_{j \geq m} c_j \right]. \quad (33)$$

This lemma will be key for  $L^2$  estimates when combined with Young's inequality:  $(a + b)^2 \leq (1 + \gamma)a^2 + (1 + \frac{1}{\gamma})b^2$  for any  $\gamma > 0$ .

**4.4.1. Discretization Error.** To pass to discrete time, we first use the classical Euler scheme for the process  $\{X_t\}$ , replacing (8) with

$$X_t^\Delta = X_{t_1}^\Delta + \mu(X_{t_1}^\Delta)(t - t_1) + \sigma(X_{t_1}^\Delta) \cdot (W_t - W_{t_1}), \quad \text{for } t \in (t_1, t_2]. \quad (34)$$

Then a standard  $L^p$ -bound for the discretization error is (Bouchard and Touzi 2004, Lemma 3.2)

$$\limsup_{\Delta t \rightarrow 0} \frac{1}{\sqrt{\Delta t}} \left\| \sup_{0 \leq t \leq T} |X_t - X_t^\Delta| \right\|_p < \infty \quad \forall p \geq 1. \quad (35)$$

Note that in some practical applications  $\{X_t\}$  can be simulated exactly, eliminating (35). In any case, assuming (34) for  $(X_t^\Delta)$ , the discretized versions of  $(Y^{k,i}, Z^{k,i}, \mathcal{M}^{k,i})$  from (25), now labelled  $(Y^{k,i,\Delta}, Z^{k,i,\Delta}, \mathcal{M}^{k,i,\Delta})$ , solve

$$\begin{cases} Y_{t_1}^{k,i,\Delta} = \mathcal{M}_{t_1}^{k,i,\Delta} \vee \left\{ \mathbb{E}[Y_{t_2}^{k,i,\Delta} | \mathcal{F}_{t_1}] + \psi_i(X_{t_1}^\Delta)\Delta t \right\}, & Y_T^{k,i,\Delta} = 0, \\ \mathcal{M}_{t_1}^{k,i,\Delta} = \max_{j \neq i} \left\{ -C_{i,j} + Y_{t_1}^{k-1,j,\Delta} \right\}, \\ Z_{t_1}^{k,i,\Delta} = \mathbb{E} \left[ Y_{t_2}^{k,i,\Delta} \cdot \frac{(W_{t_2} - W_{t_1})}{\Delta t} \middle| \mathcal{F}_{t_1} \right]. \end{cases} \quad (36)$$

The alternative

$$Y_{t_1}^{k,i,\Delta} = \max \left\{ \mathcal{M}_{t_1}^{k,i,\Delta}, \mathbb{E} \left[ Y_{t_2}^{k,i,\Delta} + \psi_i(X_{t_2}^\Delta)\Delta t \middle| \mathcal{F}_{t_1} \right] \right\} \quad (37)$$

has also been used by some authors, e.g. Bally and Pagès (2003). In the numerical analysis terminology (36) is an explicit scheme, while (37) is an implicit scheme. From a financial point of view (36) means that the payoff between today and tomorrow is certain. It is not clear which assumption is more realistic but numerically for small  $\Delta t$  the differences seem to be negligible.

The next theorem is an analogue of Theorem 7.1 in Bouchard and Touzi (2004). For completeness the proof, as well as proofs of all other theorems to follow, is provided in the Appendix.

**THEOREM 3.** *For all  $p \geq 1$ ,*

$$\limsup_{\Delta t \rightarrow 0} \frac{1}{\sqrt{\Delta t}} \sup_{0 \leq m \leq M^\#} \|Y_{m\Delta t}^{k,i,\Delta} - Y_{m\Delta t}^{k,i}\|_p < \infty. \quad (38)$$

The above error due to replacing the 'American' switching policy with a 'Bermudan' version has also been studied by Dupuis and Wang (2005). They showed that for a one-dimensional optimal stopping problem the discretized value functions in fact converge with rate  $O(\Delta t)$ , but the optimal stopping rules converge only with rate  $O(\sqrt{\Delta t})$ , as in (38).

**4.4.2. Projection Error.** For the next step we approximate the conditional expectations  $\mathbb{E}_{t_1} \triangleq \mathbb{E}[\cdot | \mathcal{F}_{t_1}]$  appearing in (36) 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  $\hat{Y}_{t_1}^{k,i}$  the approximation of  $Y_{t_1}^{k,i,\Delta}$  so that the projection is given by  $\mathcal{P}_{t_1}[\hat{Y}_{t_2}^{k,i}] = \alpha_{t_1} \cdot B_{t_1}$ , where the coefficients  $\alpha_{t_1} = \arg \min_{\alpha} \mathbb{E}[|\hat{Y}_{t_2}^{k,i} - \alpha \cdot B_{t_1}|^2]$ , are a random vector in  $\mathbb{R}^{N_B}$  depending only on  $X_{t_1}^x$ . Equation (36) becomes

$$\hat{Y}_{t_1}^{k,i} = \left( \mathcal{P}_{t_1}[\hat{Y}_{t_2}^{k,i}] + \psi_i(X_{t_1})\Delta t \right) \vee \mathcal{M}_{t_1}^{k,i}. \quad (39)$$

**THEOREM 4.** *Define  $\hat{Y}_{m\Delta t}^{k,i}$  by recursively applying (39). Then*

1.  $(\hat{Y}_{m\Delta t}^{k,i})$  is uniformly  $L^2$ -integrable in  $m$ .
2. Let  $\check{Y}_{t_1}^{k,i} \triangleq \alpha_{t_1} \cdot B_{t_1}$ . Then there exists a constant  $C$  such that

$$|\check{Y}_{t_1}^{k,i}| \leq |\alpha_{t_1}| \cdot |B_{t_1}(X_{t_1}^x)| \leq |B_{t_1}(X_{t_1}^x)| \sqrt{C(1 + \|x\|^2)}. \quad (40)$$

The bound (40) allows us to give an *a priori* bound on the regression coefficients  $\alpha_{t_1}$ . Namely, we can construct a truncation function  $T_{t_1}$  such that  $|\check{Y}_{t_1}^{k,i}| \leq T_{t_1}(X_{t_1})$  and  $\mathbb{E}[T_{t_1}(X_{t_1})^2] < \infty$ . This truncation implies that

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

We now investigate 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. (2005). It shows the relationship between the error at the  $k$ -th layer,  $|\hat{Y}_{m\Delta t}^{k,i} - Y_{m\Delta t}^{k,i}|$ , the intrinsic regression error  $\mathcal{R}_j(\mathbb{E}_m[Y_{(m+1)\Delta t}^{k,i}])$  and the error at the  $k-1$ -st layer  $|\hat{Y}_{m\Delta t}^{k-1,j} - Y_{m\Delta t}^{k-1,j}|$ .

**THEOREM 5.** *With the notation of (39) and  $Y^{k,i} \equiv Y^{k,i,\Delta}$ , there exists  $C$  such that*

$$\max_{0 \leq m \leq M^\#} \mathbb{E}[|\hat{Y}_{m\Delta t}^{k,i} - Y_{m\Delta t}^{k,i}|^2] \leq C \sum_{m=0}^{M^\#} \mathbb{E}[\mathcal{R}_m(\mathbb{E}_m[Y_{(m+1)\Delta t}^{k,i}])^2] + \frac{C}{\Delta t} \max_{0 \leq m \leq M^\#} \max_{j' \neq i} \mathbb{E}[|\hat{Y}_{m\Delta t}^{k-1,j'} - Y_{m\Delta t}^{k-1,j'}|^2]. \quad (41)$$

In a more general way we can think of  $\mathcal{P}_{t_1}$  as any approximation 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 (2004). The proof uses same techniques as for Theorem 5.

**PROPOSITION 2.** *There exists a constant  $C$  such that for any  $m, k, i$ ,*

$$\|\hat{Y}_{m\Delta t}^{k,i} - Y_{m\Delta t}^{k,i}\|_p \leq \frac{C}{\Delta t} \max_{m \leq j \leq M^\#} \left( \|(\mathbb{E}_j - \mathcal{P}_j)(\hat{Y}_{(j+1)\Delta t}^{k,i})\|_p + \max_{j' \neq i} \|\hat{Y}_{j\Delta t}^{k-1,j'} - Y_{j\Delta t}^{k-1,j'}\|_p \right). \quad (42)$$

**4.4.3. Sampling Error.** In the final step we simulate  $N$  sample paths  $\{x_{m\Delta t}^n, n = 1, \dots, N\}$  and replace the projection  $\mathcal{P}_m$  by an empirical regression. We call the resulting regression coefficients  $\alpha_{t_1}^N$  and the resulting basis functions  $B_{t_1}^n = B_{t_1}(x_{t_1}^n)$ . To control the simulated  $Y$ 's another truncation function  $T_{t_1}^n(x)$  is used such that  $T_{t_1}^n(\alpha_{t_1}^N \cdot B_{t_1}(x_{t_1}^n)) \leq 2 \cdot T_{t_1}(x_{t_1}^n)$  and  $T_{t_1}^n(\alpha_{t_1} \cdot B_{t_1}(x_{t_1}^n)) = \alpha_{t_1} \cdot B_{t_1}(x_{t_1}^n)$ .

The approximation of  $\hat{Y}^{k,i}$  along the  $n$ -th path is labelled  $\hat{Y}^{k,i,n}$  and satisfies

$$\hat{Y}_{t_1}^{k,i,n} = \left( T_{t_1}^n(\alpha_{t_1}^N \cdot B_{t_1}^n) + \psi_i(x_{t_1}^n) \Delta t \right) \vee \hat{\mathcal{M}}_{t_1}^{k,i,n}, \quad (43)$$

$$\alpha_{t_1}^N = \arg \min_{\alpha} \frac{1}{N} \sum_{n=1}^N (\hat{Y}_{t_2}^{k,i,n} - \alpha \cdot B_{t_1}^n)^2. \quad (44)$$

The truncation  $T_{t_1}^n(x)$  ensures that  $|\check{Y}_{t_1}^{k,i,n}| = |\alpha_{t_1}^N \cdot B_{t_1}^n| \leq 2 \cdot |\check{Y}_{t_1}^{k,i}|$ , which combined with (40) gives a priori control on the empirical solutions  $\hat{Y}^{k,i,n}$ . Note again that the formula (43) represents the TvR scheme (32) rather than the LS scheme (31).

Full analysis of the Monte Carlo sampling error  $\mathbb{E}[|\hat{Y}_{t_1}^{k,i} - \hat{Y}_{t_1}^{k,i,N}|^2]$ , where  $\hat{Y}_{t_1}^{k,i,N} \triangleq \frac{1}{N} \sum_{n=1}^N \hat{Y}_{t_1}^{k,i,n}$  appears to be intractable. The major difficulty is the presence of the recursively defined  $\hat{\mathcal{M}}^{k,i,n}$  which causes the sampling error to propagate in a nonlinear fashion between layers. Moreover, it is not clear how to compare the empirical  $\hat{Y}^{k,i,N}$  which is an average of pairwise maximums in (43) with the single maximum in  $\hat{Y}^{k,i}$ . Looking closely the sampling error has three components that are closely intertwined—error due to using  $\alpha_{t_1}^N$  rather than  $\alpha_{t_1}$ , error due to using  $\hat{\mathcal{M}}^{k,i,n}$  rather than  $\hat{\mathcal{M}}^{k,i}$ , and error due to potentially choosing the wrong side in the max-comparison (43), i.e. picking  $T_{t_1}^n(\alpha_{t_1}^N \cdot B_{t_1}^n) + \psi_i(x_{t_1}^n) \Delta t$  over  $\hat{\mathcal{M}}^{k,i,n}$ , when in fact  $\alpha_{t_1} \cdot B_{t_1}^n + \psi_i(x_{t_1}^n) \Delta t < \hat{\mathcal{M}}^{k,i}(x_{t_1}^n)$ .

We do not know how to isolate these error sources. This appears to be a difficult problem and existing results, even in simpler settings, are still incomplete. To summarize the current state-of-the-art in this field, let us mention the work of Clément et al. (2002) and Egloff (2005) on the Longstaff-Schwartz scheme and Bouchard and Touzi (2004), Bally and Pagès (2003) and Gobet et al. (2005) on numerical methods for backward SDEs.

For the LS scheme it has been shown that the use of optimal-stopping-time iteration in (29) is consistent, namely that for a fixed  $N_B$  the asymptotic sampling error as  $N \rightarrow \infty$  has mean zero and Gaussian distribution (Clément et al. 2002). This result was recently improved by Egloff (2005) who showed that the  $L^2$ -sampling error for a single optimal stopping problem and fixed  $\Delta t$  is  $O(\log N \cdot N^{-1})$ . However, the joint rate of convergence between the time step  $\Delta t$ , the number of basis functions  $N_B$  and the number of paths  $N$  remains unknown.

More generally, it has been proven (Bouchard and Touzi 2004, Theorem 4.1) that if one directly approximates  $\mathbb{E}_{t_1}$  with an independent Monte Carlo simulation then the resulting rate of convergence in  $L^p$ -norm is  $O(\Delta t^{-1-d/4p} \cdot N^{-2p})$ . However, the involved computational level is an order of magnitude higher. Alternatively, Gobet et al. (2005) obtain the  $O(N^{-1/2})$   $L^2$ -convergence rate when using a single set of paths but no reflection in (43). We conjecture that the sampling  $L^2$ -error is on the order  $O(\Delta t^{-1} N^{-1/2})$  for our problem as well, since this is the general convergence rate for Monte Carlo methods.

**4.4.4. Final Word On Convergence.** Summarizing the last three subsections, the total error we have is

$$\begin{aligned} \|Y_{m\Delta t}^{k,i} - \hat{Y}_{m\Delta t}^{k,i,\Delta,N}\|_2 &\leq \|Y_{m\Delta t}^{k,i} - Y_{m\Delta t}^{k,i,\Delta}\|_2 + \|Y_{m\Delta t}^{k,i,\Delta} - \hat{Y}_{m\Delta t}^{k,i,\Delta}\|_2 + \|\hat{Y}_{m\Delta t}^{k,i,\Delta} - \hat{Y}_{m\Delta t}^{k,i,\Delta,N}\|_2 \\ &\leq C \left( \sqrt{\Delta t} + \frac{(\sum_j \mathbb{E}[\mathcal{R}_{j\Delta t}(\mathbb{E}_{j\Delta t}[Y_{(j+1)\Delta t}^{1,i}]^2)])^{1/2}}{\Delta t^k} + \frac{1}{\Delta t \sqrt{N}} \right). \end{aligned} \quad (45)$$

Even without the last term which is an unproven conjecture, these convergence rates are quite conservative. We observe a general explosion in the regression errors as we cascade through the number of switches  $k$ , cf. the second term in (45). 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  $N$ . Unfortunately, Glasserman and Yu (2004) give a negative result and show that in the worst case one may need as many as  $e^{N_B}$  Monte Carlo paths to obtain a comparable sampling error. 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.

Because the LS scheme relies on stopping-time iteration, it avoids repeated conditioning and therefore regression errors are accumulated only when a *wrong* switching decision is made. This means that as long as the switching decision is correct (i.e. the correct side of the max-comparison is chosen in (43)), the pathwise value function is computed exactly (modulo discretization error). As a result, along a given path of  $\{X_t\}$ , the computation of  $J(t, x, i)$  involves only a few projection errors, rather than an error at each of  $M^\#$  steps. The number of these errors is upper bounded by the (typical) number of switching decisions made over the entire horizon. Consequently, we conjecture that the total error grows *linearly* in the maximum number of switches  $\bar{K}$ . This is corroborated by our extensive numerical experimentation. To prove this conjecture 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 expressions of the form  $\mathbb{E}[\hat{Y}_{t_2}^{k,i} \vee \hat{\mathcal{M}}_{t_1}^{k,i}]$  in terms of  $\mathbb{E}[\hat{Y}_{t_2}^{k,i}]$  and  $\mathbb{E}[\hat{\mathcal{M}}_{t_1}^{k,i}]$ , cf. (55) and proof of (54).

From our experience, the algorithm is quite stable for up to a thousand time steps,  $M^\# < 1000$ . The corresponding discretization step  $\Delta t$  should be good enough for most financial problems. The real key to successful implementation is a judicious choice of the basis functions  $(B_j)$ . After that  $N = N_p \sim 25000$  simulated paths are more than enough for the examples we tried. The empirical success of the algorithm is typical for statistical learning problems where the low but positive probability of ‘bad’ events dramatically worsens the worst case provable convergence rates (Egloff 2005) but is harmless otherwise.

To check the actual rate of convergence with respect to the size of the Monte Carlo simulation we ran Example 1 from Section 5 using 2000 – 32000 paths and tabulated the resulting means and standard deviations over 50 separate runs in Table 1. We used 400 time steps and six basis functions. We observe that the variance decreased by about 40% when we doubled the number of simulations and then stabilized, as the remaining randomness was mostly coming from the projections and not from the Monte Carlo. We also see that in this example the value function was decreasing in  $N$ , approaching the ‘true’ value of 5.931. In general, the LS scheme should provide a lower bound on  $J^k(t, x, i)$  since it finds a (sub-optimal) switching policy; however, this is often obscured by the projection and Monte Carlo errors that tend to dominate.

**Table 1** Convergence of Monte Carlo simulations for Example 1.

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

## 5. Numerical Examples.

We now give a series of numerical examples to illustrate the algorithm and the problem structure.

EXAMPLE 1. Our main example strives to give a reasonable approximation of a real gas-fired power plant that we discussed in Section 2.1. We take a 100 MW plant that can be either offline, running at half capacity of 50 MW with heat rate of 7.5 MWh/MMBtu or running at full capacity with heat rate 10 MWh/MMBTU. The maximum total annual production is therefore  $100 * 8760 = 0.876$  TW. Changing the regime of the plant requires 12 hours and 10,000 MMBtu of gas with zero output during ramping. We assume that operational decisions are made once every six hours and price the resulting tolling agreement over a period of three months (360 operational periods). For price process we take

$$\begin{aligned} dP_t &= P_t(5(50 - \log P_t) dt + 0.5 dW_t^1 + \xi_t dN_t) \\ dG_t &= G_t(2(6 - \log G_t) dt + 0.4(0.8 dW_t^1 + 0.6 dW_t^2)) \end{aligned} \quad (46)$$

where  $\xi_t \sim \exp(0.15)$  and jump intensity is  $\lambda_t = 8$  times a year. This is supposed to roughly represent energy markets in Eastern US in 2006; electricity prices are both more volatile and exhibit price spikes (only upward, with 10% increase on average), while gas prices do not. Below all values are specified on an annual basis and in millions of dollars. The profit functions are  $\psi_0(X_t) = -1$  representing fixed costs of operating a shut-down plant,  $\psi_1(X_t) = 0.438(P_t - 7.5G_t) - 1.1$  and  $\psi_2(X_t) = 0.876(P_t - 10G_t) - 1.2$  representing the spark-spread revenue plus fixed costs of operating a plant online. In line with above description, the switching costs are  $C_{01} = C_{10} = 0.01G_t$ .

EXAMPLE 2. Our second example shows some possibilities for singling out the value of different operational flexibilities. Consider a dual-mode power plant that can use either natural gas or oil. These have become popular in recent times due to extra flexibility and development of natural gas infrastructure in coastal US regions. Furthermore, natural gas prices have become less correlated with crude prices allowing for more diversification effect.

To model a dual-fuel plant, we consider a driving process  $X_t = (P_t, G_t, O_t)$  where the components are prices of electricity, gas and crude oil respectively. The state dynamics extend (46) by keeping same equations for  $(P_t)$  and  $(G_t)$  and adding

$$dO_t = O_t(2(50 - \log O_t) dt + 0.4(0.3 dW_t^1 + \sqrt{0.91} dW_t^3)), W^3 \text{ indep. of } W^1, W^2. \quad (47)$$

Observe that while oil and gas are individually correlated to power, the two input fuels are assumed to be uncorrelated between themselves.

To allow further flexibility, with either input fuel, the plant can be run at normal capacity or at full capacity which produced more power but is less efficient (due to excessive heat loss, use of older equipment, higher wear-and-tear, etc.). Thus, we assume that there are a total of five regimes available to the manager: complete shutdown  $\psi_0(X_t) \equiv 0$  as well as (all values are in millions of dollars and using annualized units):

$$\begin{aligned} \psi_1(X_t) &= 1 \cdot (P_t - 8G_t), & \psi_2(X_t) &= 1 \cdot (P_t - O_t), \\ \psi_3(X_t) &= 3 \cdot (P_t - 32G_t), & \psi_4(X_t) &= 3 \cdot (P_t - 4O_t). \end{aligned}$$

To isolate the effect of other variables, the switching costs are taken to be constant with  $C_{i,j} = 0.05$  for any  $i \neq j$ . The planning horizon is six months  $T = 0.5$  with twice-daily flexibility. While the switching costs are simplified, we assume that changing the regime takes a significant amount of time and model the delay explicitly.

In this setting, we study the effect of:

- Number of Regimes. The more operational regimes are available to the manager, the higher is the value of tolling the plant. For comparison, we check the value associated with using only a subset of the above five regimes.

- **Switching Delay.** Longer switching delays will lower the tolling value since the manager faces more risk when changing the operational regime.

**Number of Switches** Many plants have *cycling* restrictions due to engineering constraints. As a result, the manager may have absolute upper bounds on the total number of regime switches allowed (e.g. at most two each month). This will reduce value since each switch also incurs opportunity cost of reduced future flexibility.

Table 2 summarizes our findings. Removal of regimes has a strong effect. Without allowing the high-capacity oil production regime 4, the value decreases by about 9%, and without both regimes 3 and 4 which are initially 33% out-of-the-money, the expected profit declines by over 16%. Restrictions on number of switches  $k$  have an obvious and dramatic effect when  $k$  is small (nearly 20% loss of value) but less so for moderate  $k$ . It can be checked that for  $k = 7$ , the loss of value is less than 1%. Put differently, in the given model, one expects to make less than 7 switches over the next two months in nearly every case. The switching delay is less significant (at least for this example), this might have to do with the relatively small number of switches normally used, so that inability to react immediately becomes a ‘second-order’ effect.

**Table 2** Summary of results from Example 2

Setting	$J(0, X_0, 'off')$
Base Case	5.98
Regimes 0-3 only	5.33
Regimes 0-2 only	4.97
Switching delay of 1 day	6.03
Switching delay of 3 days	5.60
At most 2 switches allowed	4.86
At most 4 switches allowed	5.11

We ran the LS scheme with  $M^\# = 360$  steps and  $N = 24000$  paths. Initial prices are  $P_0 = 50, G_0 = 6, O_0 = 50$ .

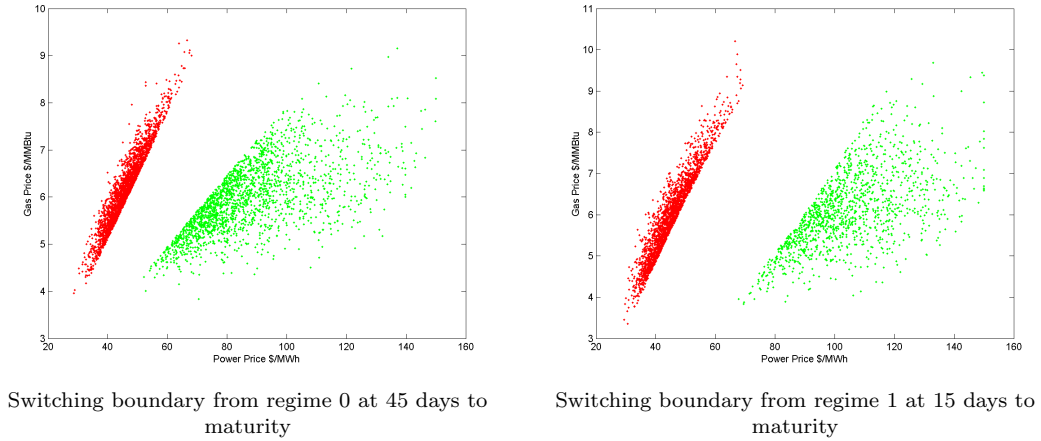
### 5.1. Approximating the Switching Boundary.

The algorithm of Section 4.3 produces a natural approximation to the switching boundary. Recall that for each  $k$  and  $i$ , the switching boundary is given by the graph  $\{(t, x_t)\}$ , such that the minimal optimal switching time of  $J^k(\cdot, i)$  is given by 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 already keep track of optimal minimal  $\tau$  in (29), 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

$$\{x_{m\Delta t}^n : n \text{ is such that } \tau^k(0, x^n, i) = m\} \quad (48)$$

defines the empirical region of switching from regime  $i$  at instant  $m\Delta t$  and we can determine the regime switched into by keeping track of  $\hat{j}^n$  from (30).

For a multidimensional  $(X_t)$  and several regimes, like in Examples 1 and 2, visualizing the switching boundaries is difficult. Figure 1 presents two slices of the switching regions for different time instances  $t$  for example 1. As can be seen the boundaries form a continuous curve in the  $(P_t, G_t)$  space, but getting a good grip on the dependence of the continuation and switching sets on  $t$  and their interrelationship for different current regimes  $i$  requires more graphing skills than we have.

**Figure 1** Optimal Exercise Boundaries for Example 1. The hysteresis band widens closer to maturity.

For a better grasp of what is going on, we illustrate the algorithm on a stylized one-dimensional example. The example uses a one-dimensional  $X_t$  (that can be interpreted as the spark spread) which can be plotted directly as a function of time. Let

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

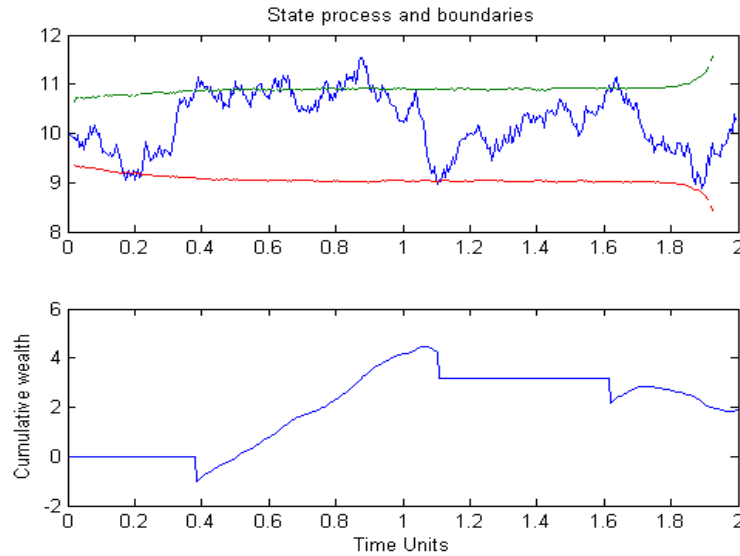
with time horizon  $T = 2$ . We have two regimes with continuous reward rates of  $\psi_0(X_t) = 0$  and  $\psi_1(X_t) = 10(X_t - 10)$ , and the switching cost between them is  $C_{0,1} = C_{1,0} = 0.3$ .

Besides illustrating the switching boundaries, we also directly show the optimal strategy for a sample path of  $(X_t)$ . Indeed, given an initial regime  $u_0$  and a realization of  $\{X_t\}$ , we keep track for each  $j \neq u_0$  of the switching sets  $Switch(u_0, j; t)$  defined in (23). At the first instant  $t$  that  $X_t \in Switch(u_0, j; t)$  for some  $j$ , we invoke our operational flexibility, switch to regime  $j$  and proceed inductively until  $T$ . The top panel of 2 shows the switching boundaries of  $J(t, x, i)$  and a realized path of  $X_t$ . 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 band where no switching takes place (note that *at-the-money* is  $x = 10$ ). Thus, if the plant is ‘off’, it will be brought online only when  $X_t$  is about 10.8. The boundaries dramatically widen close to expiration  $T$  since at that point the time decay dominates and the fixed cost is larger than any gain from switching. On the other hand, far from maturity, the exercise boundary is essentially flat as time stationarity and unlimited number of allowed switches in this example make time decay insignificant. This parallels the situation with infinite horizon where the switching boundary is constant. The bottom panel of Figure 2 shows the total cumulative wealth obtained from the policy. The discrete switching costs are indicated by drops of  $C = 0.3$  in the wealth at the switching times.

## 5.2. Alternative Computational Methods

The optimal switching problem can be solved using a variety of other numerical approaches. Roughly speaking, there are three possible strategies. First, one can directly tackle the quasi-variational formulation of (24). A number of partial differential equation (PDE) solvers can be adapted to solve the free boundary Kolmogorov equation. For example, a (a) finite-difference algorithm Wilmott et al. (1995) can be implemented. 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, since conditional expectations are transformed into weighted sums. Examples include the (b) Markov Chain approximation method of Kushner and Dupuis (2001) and the (c) Optimal Quantization technique of Bally et al. (2005).



**Figure 2** Implementing optimal policy on a one-dimensional example.

Finally, one can go back to the key representation (27) and look for other means of computing conditional expectations, see e.g. (d) kernel regression proposed by Carrière (1996).

In Ludkovski (2005), an extensive benchmarking was carried out to test the relative performance of each of the above approaches (a-d), plus the described (e) LS and (f) TvR schemes. It was observed that the LS scheme is significantly more accurate than the TvR scheme and at least on simple problems performs better than optimal quantization. Numerical solution of the QVI (24) was the most accurate but also the slowest approach. Let us stress that the benchmarking was carried out on a relatively simple 2-dimensional problem similar to Example 1. For real-life purposes, it is more important to keep in mind the robustness of different schemes and the scaling properties in the dimension of  $X_t$ . In this context, as already mentioned the QVI formulation is poor since it suffers from dimensionality curse and often necessitates use of model-specific speed-ups. Rather, we recommend the use of the LS and Quantization schemes instead. Both are simulation-based and therefore can easily incorporate many degrees of flexibility. Comparing the two, the regression Monte Carlo approach is attractive due to ease of implementation and quick setup. In contrast, quantization requires extensive pre-computation (often coupled with more sophisticated code), but offers better theoretical convergence control.

Our implementation was written in Matlab and without using any special optimization takes about five minutes to run when using 32,000 paths, 360 periods and six basis functions. The timing was done on 2.4GHz desktop. While certainly not lightning fast, we believe this is acceptable for the back-office applications we envision. Furthermore, a practical user is likely to have access to a lot more computing power and one feature of Monte Carlo algorithms is that they are naturally parallelizable.

**REMARK 2.** All the simulation-based methods above produce a lower-bound for the value of operational flexibility since they identify an approximately optimal exercise strategy. It is also possible to obtain *upper* bounds by using the duality theory developed for American options (Andersen and Broadie 2004, Haugh and Kogan 2004). Recently, this has also been extended by Meinshausen and Hambly (2004) to cover possibility of multiple exercise times. However, the implementation for our heap of optimal stopping problems is somewhat involved and is beyond the scope of this paper.

### 5.3. Comparison to Strip-of-Options Method.

We recall the currently popular method of pricing operational flexibility as a strip of spark-spread options. If for simplicity we consider just two regimes ‘off’ and ‘on’, and a set of “exercise times”  $T_m = m\Delta t$ , this means that the overall value of running the power plant is approximated by

$$V(0, x; \Delta t) = \mathbb{E}^x \left[ \sum_{m=0}^{M^\sharp} (P_{m\Delta t} - \overline{HR} \cdot G_{m\Delta t} - K)^+ \cdot \Delta t \right], \quad (P_0, G_0) = x. \quad (49)$$

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  $C = \max_{i,j} C_{i,j}$ . In the limit of vanishing switching costs  $C \rightarrow 0$  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  $\Delta t$  and current time  $t_1$ . Clearly,  $V(t_1, x)$  is the maximum value that one can extract from the plant since there are no operational constraints, so trivially  $V(t_1, x) \geq J(t_1, x, i)$ . By a basic conditioning,

$$V(t_1, x) = \mathbb{E}[V(t_2, X_{t_2}^{t_1, x}) | X_{t_1} = x] + (P_{t_1} - \overline{HR} \cdot G_{t_1} - K)^+ \cdot \Delta t. \quad (50)$$

Next observe that for any  $x$  there exists mode  $i$  such that  $\psi_i(x) = (P_{t_1} - \overline{HR} \cdot G_{t_1} - K)^+$ . For this mode  $i$ ,  $J(t_1, x, i) \geq \mathbb{E}[J(t_2, X_{t_2}^{t_1, x}, i) | X_{t_1} = x] + \psi_i(x)\Delta t$ , and combining with (50),

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

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

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

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

With the strip of options approach even when  $C_{i,j}$ ’s are significant, the switching costs do not enter into the decision process so that many exercises are made to capture relatively small gains. A possible fix is to compute  $V(t, x)$  and then estimate the corresponding number of switches. This can be done by picking an approximate set of switching boundaries and counting the number of times  $(X_t)$  will cross it. The last part can be done via simulation. The method can provide rough answers, but of course finding a reasonable set of switching boundaries in a multidimensional setup is entirely non-obvious and is in fact as hard as the original problem.

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  $J^k$ .

## 6. Extensions.

### 6.1. 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 mining site or an 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 Schwartz (1985). However, for realistic financial planning the horizon should be finite. Most 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.

Management of exhaustible resources is easily convertible into our recursive optimal stopping framework. Let us postulate again that the mine can be operated in a discrete number of regimes with production rates  $a_i$ . Let  $I_t$  denote the current level of resources left (thus  $I_t$  is a non-increasing process). Given initial known inventory of  $I_0$  the objective is to maximize the total profit so that the value function  $J(t, x, c, i)$  satisfies

$$\begin{cases} J(t, x, c, i) = \sup_{u \in \mathcal{U}(t)} \mathbb{E} \left[ \int_t^T \psi_{u_s}(s, X_s) ds - \sum_{\tau_k < T} C_{u_{\tau_k^-}, u_{\tau_k}} \middle| X_t = x, I_t = c, u_t = i \right] \\ J(t, x, 0, i) \equiv 0, \quad I_t \geq 0. \end{cases} \quad (51)$$

The boundary condition models resource depletion after which the mine must be mothballed.

The new complexity in (51) arises due to the inventory  $I_t$ , which depends on past history and specifically the policy chosen over  $[0, t]$ . However, we observe that  $(I_t)$  is degenerate in the sense that it can be deterministically computed from knowledge of initial inventory level  $I_0$  and scheduling policy  $u$ . In particular, if  $\tau$  denotes the time of next switching decision then (51) can be re-written as

$$J(t, x, c, i) = \sup_{\tau \in \mathcal{S}_t, j \neq i} \mathbb{E} \left[ \int_t^\tau \psi_i(s, X_s) ds + J(\tau, X_\tau, I_\tau, j) - C_{i,j} \middle| X_t = x, I_t = c \right], \quad (52)$$

and  $I_\tau = c - a_i(\tau - t)$ , allowing us to effectively eliminate  $(I_t)$ . As a result, given  $\Delta t$  we can construct a finite grid  $I_t \in \{m\Delta c: 0 \leq m \leq N_c\}$ , such that the remaining inventory following any policy  $u$  always lies on the grid points,  $\exists m'$ , s.t.  $I_{m\Delta t} = m'\Delta c$ . After that, the framework of Section 3 can be straightforwardly re-used, see Ludkovski (2005) for full details.

**EXAMPLE 3.** For comparison we take up the example of mine management in Brennan and Schwartz (1985). Consider a copper mine with finite inventory worth 15 years of production. The mine can be either operated at a fixed rate or kept idle and the license expires after 99 years. The price process is a one-dimensional geometric Brownian motion

$$dX_t = 0.01X_t dt + \sqrt{0.08}X_t dW_t,$$

and the payoff rates in annualized millions of dollars are

$$\psi_0(X_t) = -0.5 \quad \text{and} \quad \psi_1(X_t) = 10(X_t - 0.5) - 5(X_t - 0.5)^+.$$

The switching cost is  $C_{0,1} = C_{1,0} = 0.2$  and there is a discount rate of  $r = 4\%$ . In addition, the mine operator has the option of complete abandonment, shutting down all operations and exiting the business. Such extra flexibility provides insurance against really low prices and has been also studied by Zervos (2003). In our framework, we implement it by directly imposing the constraint  $J(m\Delta t, x, i) \geq 0$  during Dynamic Programming.

The original paper of (Brennan and Schwartz 1985, p. 343) solves the problem by using a PDE solver applied to the corresponding quasi-variational inequality. In contrast, we utilize our extended regression Monte Carlo scheme described in the last paragraph. We took 600 time steps, 10000 paths and 75 gridpoints in the inventory  $I$ -space. The results are summarized in Table 3. We find that our simulation-based solver (labelled CL '06) has errors of less than 2% everywhere and is an order faster than the PDE method (which we reproduced through a basic FD method), already showing significant savings for this toy example.

**Table 3** Comparison of pricing methods for the mining example of Brennan and Schwartz (1985), varying the initial copper price  $P_0$

Method/ $P_0$	0.3	0.4	0.5	0.6	0.7	0.8
BS '85	1.45	4.35	8.11	12.49	17.38	22.68
CL '06	1.35	4.41	8.15	12.44	17.52	22.41

The initial inventory is  $I_0 = 150$  million pounds and the values are in millions of dollars.

## 6.2. Plant Outages.

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 expected benefit of being online is reduced. For simplicity let us assume that emergency maintenance shutdowns are completely unpredictable and occur independently with a constant intensity rate  $\lambda$ . Letting  $\tilde{T}_k$  represent the random time of the  $k$ -th emergency, we have  $\mathbb{P}[\tilde{T}_k \geq s + \tilde{T}_{k-1}] = e^{-\lambda s}$ . Since the times between outages have the memoryless exponential distribution,  $\lambda$  acts like an additional discount factor and the recursive construction (11) becomes

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

The first term above represents the probability of an outage and a forced switch to mode ‘off’ before  $\tau$ . Related work on investment with forced exit has been done by Wang (2005). Numerically, in discrete time the benefit of no-switch at instant  $t$  is simply changed to  $\mathbb{E}[(1 - \lambda\Delta t)J^k(t + \Delta t, X_{t+\Delta t}, i) + \lambda\Delta t(J^{k-1}(t + \Delta t, X_{t+\Delta t}, 0) - C_{i,0}) | \mathcal{F}_t] + \psi_i(t, X_t)\Delta t$ .

## 6.3. Further Features.

From the point of view of realistic implementation, it is clear that the models presented so far are dramatic simplifications. Accordingly, let us briefly discuss what other features need to be added for practical use. To begin, a satisfactory description of electricity and/or gas prices is likely to use several stochastic factors or a regime-switching model. Since our model is stated for generic  $\{X_t\}$ , there is no difficulty in extending the dimension. Various forms of non-stationarity, including seasonality and time-dependent coefficients are also certainly needed. These are similarly straightforward to incorporate, as long as we have a method for simulating the forward  $\{X_t\}$  paths. Clearly, any agent wishing to price an  $\{X_t\}$ -payoff should already have handy some scheme for such simulations. Extra operational constraints can be dealt with by modifying the set of acceptable policies  $\mathcal{U}(t)$  in (6). For instance, one can add planned maintenance shutdowns or requirements for running the plant at specific times.

Price impact is another important feature. A simple version can postulate that the agent moves the market by a fixed amount (either directly or proportionally). Such an effect can be directly incorporated into the actual profit functions  $\psi_m$ . A more sophisticated effect can be obtained by relying on supply-demand equilibrium. Suppose that the supply is fixed at level  $S$ . The demand  $D_t$  on the other hand is stochastic and fluctuates such that the clearing price is determined by a demand-supply equilibrium. Then  $P_t = g(D_t, S + \xi_t)$  where  $\xi_t$  is the supply generated by the plant itself. Observe that after adding  $D_t$  as a state variable, we can still write  $\psi_m(P_t, G_t, D_t)$  preserving a Markov state space. An important application of this situation concerns *guaranteed serving loads*,

a contract where the power company agrees to supply as much electricity as demanded by the retail customers and exposing itself to volume risk.

Finally, it is also possible to imagine that actions of the agent influence the *parameters* of the price processes, so that e.g. the mean-reversion level of  $P_t$  depends on the current regime. Such situation can also be dealt with straightforwardly. Notice that we just need to simulate  $P_t$  under each regime and then use the Longstaff-Schwartz scheme over each set of paths. The backward recursion algorithm then proceeds by resetting the value functions using the approximate conditional expectation  $\hat{E}$  after each switching decision. Such a mixed TvR-LS scheme has been generally described by Egloff (2005).

Some potentially useful features are unsuitable for our framework. These include any ‘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. Moreover,  $\psi_i(s, X_s)$  may depend on the previous regime (e.g. the heat rate will be larger when the plant is cold compared to when it has been running and is warm). In principle, such situations can be incorporated by enlarging the state space to preserve the Markovianity of  $\{X_t\}$ . However, on a practical level this is likely to be infeasible due to explosion of computational complexity.

#### 6.4. Conclusion.

This paper investigated the numerical solution of operational flexibility problems that arise in management of energy assets. Our key proposal is a stochastic control framework that allows natural merging between the operational constraints and the dynamic nature of the problem. Moreover our framework is tractable since it is equivalent to a cascade of optimal stopping problems. Those can be solved by numerically computing the corresponding Snell envelopes via simulation and backward recursion. 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 current practice of strip-of-spread-options that ignores the operational constraints of tolling agreements.

The techniques described are applicable to many other energy-related problems. In particular, gas storage and power supply guarantees are important practical challenges on which our method can shed new light. Full analysis of these two settings is left to the companion paper (Carmona and Ludkovski 2005). For 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 Section 3.3.

#### Appendix. Proofs for Section 4.4.

**Proof of Theorem 3.** As in all subsequent proofs we will proceed by induction on  $k$ . Let  $\mathcal{S}^\Delta(m\Delta t) = \{j\Delta t : m \leq j \leq M^\#\}$  be the set of discretized stopping times after  $m\Delta t$  and define the auxiliary process

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

By a standard argument under Assumption 2 (see e.g. Bally and Pagès (2003), Lemma 3 and 4),  $\max_{0 \leq m \leq M^\#} \|R_{m\Delta t}^{k,i} - Y_{m\Delta t}^{k,i}\|_p \leq C\Delta t$ . To estimate the difference of  $R_{m\Delta t}^{k,i}$  and  $Y_{m\Delta t}^{k,i,\Delta}$  which use

different barriers and different  $X$ 's, apply  $|(X_1 \vee Y_1) - (X_2 \vee Y_2)| \leq |X_1 - X_2| \vee |Y_1 - Y_2|$  on  $|\mathcal{M}_\tau^{k,i} - \mathcal{M}_\tau^{k,i,\Delta}|$  to get

$$\begin{aligned} |R_{m\Delta t}^{k,i} - Y_{m\Delta t}^{k,i,\Delta}| &\leq \operatorname{ess\,sup}_{\tau \in S^\Delta(m\Delta t)} \mathbb{E}_m \left[ |\mathcal{M}_\tau^{k,i} - \mathcal{M}_\tau^{k,i,\Delta}| + \Delta t \sum_{j=m}^{M^\#} \mathbb{1}_{\tau > j\Delta t} \cdot |\psi_i(X_{j\Delta t}^\Delta) - \psi_i(X_{j\Delta t})| \right] \\ &\leq \mathbb{E}_m \left[ \max_{m \leq j \leq M^\#} \max_{j' \neq i} |Y_{j\Delta t}^{k-1,j'} - Y_{j\Delta t}^{k-1,j',\Delta}| + C |X_{j\Delta t}^\Delta - X_{j\Delta t}| \right] \end{aligned}$$

or  $\|R_{m\Delta t}^{k,i} - Y_{m\Delta t}^{k,i,\Delta}\|_p \leq C\sqrt{\Delta t} + \left\| \max_{0 \leq j \leq M^\#} \max_{j' \neq i} |Y_{j\Delta t}^{k-1,j'} - Y_{j\Delta t}^{k-1,j',\Delta}| \right\|_p$  by (35). Combining the two bounds we obtain,

$$\left\| \max_{0 \leq m \leq M^\#} |Y_{m\Delta t}^{k,i,\Delta} - Y_{m\Delta t}^{k,i}| \right\|_p \leq C\sqrt{\Delta t} + \left\| \max_{0 \leq m \leq M^\#} \max_{j' \neq i} |Y_{m\Delta t}^{k-1,j'} - Y_{m\Delta t}^{k-1,j',\Delta}| \right\|_p$$

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

**Proof of Theorem 4.** We first check that for a fixed  $k$ ,  $\hat{Y}_{m\Delta t}^{k,i}$  is uniformly  $L^2$ -integrable in time. Using

$$\mathbb{E}|(X_1 \vee Y_1) - (X_2 \vee Y_2)|^2 \leq (1 + \gamma)\mathbb{E}|X_1 - X_2|^2 \vee (1 + \frac{1}{\gamma})\mathbb{E}|Y_1 - Y_2|^2 \quad (53)$$

for any  $\gamma > 0$ , as well as  $\mathcal{P}_{t_1}$  being an  $L^2$ -contraction and Young's inequality we obtain

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

so that by Gronwall's lemma

$$\max_{0 \leq m \leq M^\#} \mathbb{E}[|\hat{Y}_{m\Delta t}^{k,i}|^2] \leq C(1 + \|x\|^2) + \frac{1}{\Delta t} \max_{j \neq i} \max_{0 \leq m \leq M^\#} \mathbb{E}[|\hat{Y}_{m\Delta t}^{k-1,j}|^2]. \quad (54)$$

Completing the induction on  $k$ ,  $\max_{0 \leq m \leq M^\#} \mathbb{E}[|\hat{Y}_{m\Delta t}^{k,i}|^2] \leq C(\Delta t)^{-k}(1 + \|x\|^2)$ . This provides a priori estimates on the regression result  $\hat{Y}_{t_1}$ . By (54) and orthonormality of  $B_{t_1}$  we have  $\mathbb{E}|\hat{Y}_{t_1}^{k,i}|^2 = \mathbb{E}[|\alpha_{t_1}|^2 |B_{t_1}|^2] = |\alpha_{t_1}|^2 \leq C(1 + \|x\|^2)$  and (40) follows.  $\square$

**Proof of Theorem 5.** 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  $\eta_{t_1}^{k,i} \triangleq \mathbb{E} \left[ |\hat{Y}_{t_1}^{k,i} - Y_{t_1}^{k,i}|^2 \right]$ . We will show that

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

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

$$\begin{aligned} \eta_{t_1}^{k,i} &= \mathbb{E} \left[ \left\{ \mathcal{P}_{t_1}(\hat{Y}_{t_2}^{k,i}) + \psi_i(t_1, X_{t_1})\Delta t \vee \hat{\mathcal{M}}_{t_1}^{k,i} \right\} - \left\{ \mathbb{E}_{t_1}(Y_{t_2}^{k,i}) + \psi_i(t_1, X_{t_1})\Delta t \vee \mathcal{M}_{t_1}^{k,i} \right\} \right]^2 \\ &\leq \mathbb{E} \left[ |\mathcal{P}_{t_1}(\hat{Y}_{t_2}^{k,i}) - \mathbb{E}_{t_1}(Y_{t_2}^{k,i})|^2 \vee |\hat{\mathcal{M}}_{t_1}^{k,i} - \mathcal{M}_{t_1}^{k,i}|^2 \right] \end{aligned} \quad (55)$$

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

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

by the orthogonality of the remainder  $\mathcal{R}_{t_1}$ . Substituting into (55) and using (53)

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

which immediately implies (41).  $\square$

### Acknowledgments

We thank Monique Jeanblanc, Bernard Lapeyre and Nizar Touzi for enlightening discussions on optimal switching, the Longstaff-Schwartz algorithm and BSDE's.

### References

- Andersen, L., M. Broadie. 2004. A primal-dual simulation algorithm for pricing multi-dimensional American options. *Management Sci.* **50**(9) 1222–1234.
- Bally, V., G. Pagès. 2003. Error analysis of the optimal quantization algorithm for obstacle problems. *Stochastic Process. Appl.* **106**(1) 1–40.
- Bally, V., G. Pagès, J. Printems. 2005. A quantization tree method for pricing and hedging multidimensional American options. *Math. Finance* **15**(1) 119–168.
- Bar-Ilan, A., A. Sulem, A. Zanello. 2002. Time-to-build and capacity choice. *J. Econom. Dynam. Control* **26**(1) 69–98.
- Bensoussan, A., J-L. Lions. 1984. *Impulse Control and Quasi-Variational Inequalities*. Gauthier-Villars, Paris.
- Bouchard, B., N. Touzi. 2004. Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Process. Appl.* **111**(2) 175–206.
- Brekke, K.A., B. Øksendal. 1994. Optimal switching in an economic activity under uncertainty. *SIAM J. Control Optim.* **32**(4) 1021–1036.
- Brennan, M., E. Schwartz. 1985. Evaluating natural resource investments. *J. Business* **58** 135–157.
- Carmona, R., V. Durrleman. 2003. Pricing and hedging spread options. *SIAM Rev.* **45**(4) 627–685.
- Carmona, R., M. Ludkovski. 2005. An impulse control approach to gas storage and power supply guarantees Submitted.
- Carmona, R., N. Touzi. 2005. Optimal multiple stopping and valuation of swing options.
- Carrière, J. F. 1996. Valuation of the early-exercise price for options using simulations and nonparametric regression. *Insurance : Math. Econom.* **19** 19–30.
- Chevance, D. 1997. Numerical methods for backward stochastic differential equations. L.C.G. Rogers, D. Talay, eds., *Numerical methods in finance*. Cambridge University Press, 232–244.
- Clément, E., D. Lamberton, P. Protter. 2002. An analysis of a least squares regression algorithm for American option pricing. *Finance Stoch.* **6** 449–471.
- Deng, S.J., Z. Xia. 2005. Pricing and hedging electric supply contracts: a case with tolling agreements. Preprint.
- Dixit, A. 1989. Entry and exit decisions under uncertainty. *J. Polit. Economy* **97**(3) 620–638.
- Duan, J.C., S. Pliska. 2004. Option valuation with co-integrated asset prices. *J. Econom. Dyn. Control* **28** 727–754.
- Dupuis, P., H. Wang. 2005. On the convergence from discrete to continuous time in an optimal stopping problem. *Ann. Appl. Probab.* **15** 1339–1366.

- Egloff, D. 2005. Monte Carlo algorithms for optimal stopping and statistical learning. *Ann. Appl. Probab.* **15**(2) 1396–1432.
- El-Karoui, N. 1981. *Ecole d'été de Saint-Flour, Lecture Notes in Mathematics*, vol. 876, chap. Les aspects probabilistes du contrôle stochastique. Springer, Berlin, 73–258.
- El-Karoui, N., C. Kapoudjian, E. Pardoux, S. Peng, M.C. Quenez. 1997. Reflected solutions of backward SDE's and related obstacle problems for PDE's. *Ann. Probab.* **25** 702–737.
- Eydeland, A., K. Wolyniec. 2003. *Energy and Power Risk Management: New Developments in Modeling, Pricing and Hedging*. John Wiley&Sons, Hoboken, NJ.
- Glasserman, P., B. Yu. 2004. Number of paths versus number of basis functions in American option pricing. *Ann. Appl. Probab.* **14**(4) 2090–2119.
- Gobet, E., J.P. Lemor, X. Warin. 2005. A regression-based Monte-Carlo method to solve backward stochastic differential equations. *Ann. Appl. Probab.* **15**(3) 2172–2002.
- Guo, X., H. Pham. 2005. Optimal partially reversible investment with entry decision and general production function. *Stochastic Process. Appl.* **115** 705–736.
- Hamadène, S., M. Jeanblanc. 2004. On the starting and stopping problem: application in reversible investments Preprint.
- Hamadène, S., J.P. Lepeltier. 2000. Reflected BSDEs and mixed game problem. *Stochastic Process. Appl.* **85**(2) 177–188.
- Haug, M., L. Kogan. 2004. Pricing American options: A duality approach. *Oper. Res.* **52**(2) 258–270.
- Jeanblanc, M., A. Shiryaev. 1995. Optimization of the flow of dividends (in russian). *Russian Math Surveys* **50** 257–277.
- Kushner, H.J., P. Dupuis. 2001. *Numerical methods for stochastic control problems in continuous time, Applications of Mathematics*, vol. 24. 2nd ed. Springer-Verlag, New York.
- Lepeltier, J.P., B. Marchal. 1984. Théorie générale du contrôle impulsif markovien. *SIAM J. Control Optim.* **22**(4) 645–665.
- Longstaff, F.A., E.S. Schwartz. 2001. Valuing American options by simulations: a simple least squares approach. *Review Finan. Stud.* **14** 113–148.
- Ludkovski, M. 2005. Optimal switching with application to energy tolling agreements. Ph.D. thesis, Princeton University.
- Ludkovski, M. 2006. Financial hedging of operational flexibility. Working Paper, University of Michigan.
- Meinshausen, N., B.M. Hambly. 2004. Monte Carlo methods for the valuation of multiple exercise options. *Math. Finance* **14** 557–583.
- Øksendal, B., A. Sulem. 2005. *Applied stochastic control of jump diffusions*. Springer-Verlag, Berlin.
- Pham, H. 2005. On the smooth-fit property for one dimensional optimal switching problem. *Séminaire de Probabilités, vol XL*. To appear.
- Pham, H., V. Ly Vath. 2005. Explicit solution to an optimal switching problem in the two-regime case Working Paper.
- Protter, P. 2004. *Stochastic integration and differential equations, Applications of Mathematics*, vol. 21. Springer-Verlag, Berlin. 2nd Edition.
- Tsitsiklis, J. N., B. Van Roy. 2001. Regression methods for pricing complex American-style options. *IEEE Trans. Neural Netw.* **12**(4) 694–703.
- Wang, H. 2005. A sequential entry problem with forced exits. *Math. Oper. Res.* **30** 501–520.
- Wilmott, P., S. Howison, J. Dewynne. 1995. *The mathematics of financial derivatives*. Cambridge University Press, Cambridge. A student introduction.
- Yushkevich, A. 2001. Optimal switching problem for countable Markov chains: average reward criterion. *Math. Methods Oper. Res.* **53**(1) 1–24.
- Zervos, M. 2003. A problem of sequential entry and exit decisions combined with discretionary stopping. *SIAM J. Control Optim.* **42**(2) 397–421.