

PRICING ASSET SCHEDULING FLEXIBILITY USING OPTIMAL SWITCHING

RENÉ CARMONA AND MICHAEL LUDKOVSKI

ABSTRACT. 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 computational examples. We benchmark and compare our scheme to alternative numerical methods.

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 to 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. Our method originates with Brennan and Schwartz (1985), but only recently has it received a rigorous analysis.

The existing mathematical finance literature related to the optimal switching problem (Dixit 1989, Brekke and Øksendal 1994, Zervos 2003, Guo and Pham 2005, Pham and Ly Vath 2007, Bayraktar and Egami 2007) 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

Date: First Version: July 2005. This Version: November 26, 2007.

2000 Mathematics Subject Classification. Primary: 93E20, 65C05; Secondary: 60G40, 60H35.

Key words and phrases. optimal switching, Monte Carlo, operational flexibility, impulse control, Snell envelope.

problems. Concentrating on contracts with fixed expiry dates leads to a more difficult problem but stays closer to reality. 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 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.

Remark 1. After this work was completed, we learned of the existence of a parallel work by Deng and Xia (2005). While addressing similar issues, Deng and Xia (2005) is restricted to the Tsitsiklis and van Roy form (cf. (34)) of the regression algorithm for the dynamic programming equations. Moreover, the authors do not provide the rigorous convergence analysis of the numerical scheme or the proof for the identification of the continuation and switching regions.

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 showcase 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.

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 fuel (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 (2007) 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; the latter have been intensively studied in classical option pricing, see for instance an overview in 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. 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 of 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 prices 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, then when it runs its turbines the electricity price P_t is expected to decrease due to increased supply. For notational clarity we omit this situation for the main model and return to it in Section 6.3. 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) revenue rates.

Changing an output level is costly, requiring extra fuel and various overhead costs. We denote by $C_{i,j}$ the switching costs from regime i to regime 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.

Assumption 1. *The switching costs are strictly positive with $C_{i,j} > \epsilon_C > 0$, for all $i \neq j$ and some $\epsilon_C > 0$. Also, $C_{i,i} = 0$, and $(C_{i,j})$ satisfies the triangle inequality*

$$C_{i,j} \leq C_{i,k} + C_{k,j}, \quad \text{for any } i, j, k. \quad (2)$$

The latter condition 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 operator 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 start-up and shut-down 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

$$u = ((\xi_1, \tau_1), (\xi_2, \tau_2), \dots), \quad \text{with } \xi_\ell \in \mathbb{Z}_M \triangleq \{0, \dots, M-1\}, \text{ and } 0 \leq \tau_{\ell-1} < \tau_\ell < T, \quad (3)$$

where ξ_ℓ are the successive modes chosen, and τ_ℓ are the (\mathcal{F}_t^X) -stopping) switching times. Due to Assumption 1 regarding subadditive costs, multiple simultaneous switches with $\tau_\ell = \tau_{\ell+1}$ are suboptimal. From a control-theoretic perspective, one should think of u as an \mathcal{F}^X -adapted right-continuous piecewise-constant process where $u_t = \sum_{\tau_\ell < T} \xi_\ell \mathbb{1}_{[\tau_\ell, \tau_{\ell+1})}(t)$ denotes the operating mode at time t .

The total reward for such control u and scenario $\omega \in \Omega$ is

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

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 τ_ℓ . The payoff rates ψ_m 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 $\psi_m(t, X_t) = e^{-rt}u(f_m(X_t))$, where u is a concave utility function and $f_m(X_t)$ is the (possibly negative) nominal income stream from the asset.

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

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

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 an (ϵ -)optimal switching policy u^ϵ for (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

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

which is the difference between expected profit with the flexibility built-in and the expected profit in the best 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, see Example 3 in Section 5.

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 and does not cause jumps in (X_t) , it is called *optimal switching*. In the general form considered here (observe that we have minimal structure imposed on (X_t) , ψ_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^ϵ .

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 τ_ℓ and choice/size of project to start ξ_ℓ , similar to the structure of (3). An 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 τ_ℓ and dividend amounts ξ_ℓ , 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 (2007) 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 (2007), see also Bayraktar and Egami (2007). 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 (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 she 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 and derive extra benefits that are similar to the time premium for American options. See Section 5.1 for further discussion of these effects.

3. SOLUTION METHODOLOGY.

In this section we describe the stochastic framework for the optimization problem (5) and characterize its solution in terms of a limit of iterated optimal stopping problems. We then describe connections to the variational and backward SDE formulations.

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 $(X_t) = (X_t^1, \dots, X_t^d)$ whose dynamics are given by

$$dX_t = \mu(X_t) dt + \sigma(X_t) \cdot dW_t + \int_{\mathbb{R}^d} \iota(X_{t-}, y) N(dt, dy), \quad (7)$$

where $\{W_t\}$ is a standard e -dimensional Wiener process on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and N is a Poisson random measure on $[0, \infty) \times \mathbb{R}^d$, independent of $\{W_t\}$. The compensator of N is postulated to be of the form $\lambda(X_{t-}) dt \cdot \nu(dy)$, where λ is the state-dependent jump intensity and ν is the jump distribution. The filtration \mathbb{F} satisfies the *usual* conditions, with \mathcal{F}_0 being trivial. Each of μ , σ , ι and (λ, ν) 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) (7) is non-degenerate in the sense that it has a unique strong solution and a positive-definite diffusion matrix, $x \cdot \sigma^t(x) \sigma(x) \cdot x^T > \epsilon > 0$, $\forall x \in \mathbb{R}^d$. We shall also use the notation

$$\mathcal{S}_T^p \triangleq \left\{ Z: \mathcal{F}_t^X \text{ - adapted, } \mathbb{E} \left[\sup_{t \in [0, T]} |Z_t|^p \right] < \infty \right\}, \quad p \geq 1,$$

and make the following standing assumption:

Assumption 2. 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 satisfied for example if each component of (X_t) is sup-square integrable $X^n \in \mathcal{S}_T^2$, $n = 1, \dots, d$, and all the rewards are of linear growth uniformly in time, $|\psi_m(t, x)| < C(1 + \|x\|)$, where $\|x\| = \sum_{i=1}^d x_i^2$ denotes the Euclidean norm on \mathbb{R}^d .

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, $\mathbb{P}[\tau_k < T \ \forall k \geq 0] = 0$. In the usual case where the expected maximum gain is bounded, $\mathbb{E}[\int_0^T \max_i |\psi_i(s, X_s^x)| ds] < \infty$, this restriction is superfluous 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

$$dX_t^n = X_t^n [\kappa^n (\theta^n - \nu^n \lambda^n \log X_t^n) dt + \Sigma^n \cdot dW_t + d(\sum_{i=1}^{N_t^n} (e^{\xi_i^n} - 1))], \quad n = 1, \dots, d, \quad (8)$$

$$\text{or } d(\log X_t^n) = \kappa^n (\theta^n - \nu^n \lambda^n - \frac{\sigma_n^2}{2\kappa^n} - \log X_t^n) dt + \Sigma^n \cdot dW_t + \xi^n dN_t, \quad X_0^n = x^n,$$

where $\{W_t\}$ is a d -dimensional standard Wiener process, N_t^n are independent Poisson processes with intensity λ^n , $\sigma_n^2 = (\Sigma^{n,1})^2 + \dots + (\Sigma^{n,d})^2$, $\Sigma = [\Sigma^{n,p}]_{n,p=1,\dots,d} \in \mathbb{R}^{d \times d}$ is a constant non-degenerate volatility matrix, and $\xi_i^n \sim \text{Exp}(1/\nu^n)$. The last statement means that the n -th component X^n experiences jumps of mean size ν^n -percent that occur at constant rate λ^n . Notice that the dependence between the components of (X_t) is only through the matrix Σ (dependence between the jump times can also be added). The model is meant to reflected the widely-documented price spikes and mean-reversion observed in commodity markets. The attractiveness of (8) 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 2 is automatically satisfied for any reward of polynomial growth.

There is still an ongoing debate about appropriate assumptions about electricity and gas 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 (Eydeland and Wolyniec 2003) have suggested that the spark spread $P_t - \overline{H\overline{R}} \cdot G_t$ 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

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

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

$$I^k(t, x, i) = \sup_{u \in \mathcal{U}^k(t)} \mathbb{E} \left[\int_t^T \psi_{u_s}(s, X_s^{t,x}) ds - \sum_{t \leq \tau_\ell < T} C_{u_{\tau_\ell-}, u_{\tau_\ell}} \Big| X_t^{t,x} = x, u_t = i \right]. \quad (9)$$

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 τ , which maximizes the initial payoff until τ , plus the value function at τ 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 ν , let $\mathcal{S}_\nu \triangleq \{\tau \leq T : \mathbb{F}\text{-stopping time s.t. } \nu \leq \tau \text{ a.s.}\}$ be the set of all stopping times after ν . Thus, $\mathcal{S} \equiv \mathcal{S}_0$ is the set of all \mathbb{F} -stopping times bounded by T . Define a functional operator \mathcal{M} whose action on test function w is given by

$$\mathcal{M}w(t, x, i) \triangleq \max_{j \neq i} \left\{ -C_{i,j} + w(t, x, j) \right\}. \quad (10)$$

We iteratively define $J^k(t, x, i)$ for $t \in [0, T]$, $i \in \mathbb{Z}_M$, by

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

Thus, 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. If there are only two regimes like in Hamadène and Jeanblanc (2007), the maximum in (10) 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 recall the general theory of Snell envelopes, summarized in the following proposition (note that it does not make any Markovian assumptions).

Proposition 1. (*Hamadène and Lepeltier 2000*) *Let Z be an \mathbb{R} -valued rcll 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*

$$Y(\nu) \triangleq \operatorname{ess\,sup}_{\tau \in \mathcal{S}_\nu} \mathbb{E} \left[Z_\tau | \mathcal{F}_\nu \right]. \quad (12)$$

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 ν . Y^r is called the Snell envelope of Z .*
- (2) *The minimal optimal stopping time τ^* 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 \mathcal{F} -supermartingale on $[0, T]$ and a \mathcal{F} -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, we have that $J^0(t, X_t^{0,x}, i)$ satisfies all the regularity assumptions of Proposition 1. Pick an initial value $X_0^{0,x} = x$ and let $\Psi_t^i = \int_0^t \psi_i(s, X_s^{0,x}) ds$, $Z_t^{k,i} = \int_0^t \psi_i(s, X_s^{0,x}) ds + \mathcal{M}J^{k-1}(t, X_t^{0,x}, i)$. Assume by an induction hypothesis that $Z^{k,i}$ satisfies the assumptions of Proposition 1. Then $Y^{k,i}$, the Snell envelope of $Z^{k,i}$, solves $Y_t^{k,i} = \operatorname{ess\,sup}_{\tau \in \mathcal{S}_t} \mathbb{E}[Z_\tau^{k,i} | \mathcal{F}_t]$ and comparing with (11) we see that

$$J^k(t, X_t^{0,x}, i) = \operatorname{ess\,sup}_{\tau \in \mathcal{S}_t} \mathbb{E} \left[\int_t^\tau \psi_i(s, X_s^{0,x}) ds + \mathcal{M}J^{k-1}(\tau, X_\tau^{0,x}, i) \middle| \mathcal{F}_t \right] = Y_t^{k,i} - \Psi_t^i. \quad (13)$$

Accordingly, $J^k(t, X_t^{0,x}, i)$ is a square-integrable \mathcal{F} -supermartingale, and if X satisfies item 4 of Proposition 1, is continuous. This identification also shows that the minimal optimal stopping time τ_k^* corresponding to the Snell envelope defined by $J^k(t, X^{t,x}, i)$ is simply

$$\tau_k^* = \inf\{s \geq t: J^k(s, X_s^{t,x}, i) = \mathcal{M}J^{k-1}(s, X_s^{t,x}, i)\} \wedge T. \quad (14)$$

Summarizing, the induction argument implies that for all k , $\{J^k(t, X_t^{0,x}, i)\}_t$ is a measurable, right-continuous with left limits process of class [D]. Armed with these facts we next show that the recursive definition is equivalent to the original restricted problem in (9).

Theorem 2. J^k is equal to the value function for the optimal switching problem with at most k switches allowed, I^k .

Proof. The theorem states that a ‘global’ optimization using a fixed maximum of k switches in (9) 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, $I^0 = J^0$, since the sum appearing in the right hand side of definition (9) is not present for $k = 0$. Take $k \geq 1$ and assume that $I^{k-1} = J^{k-1}$. 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

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

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 $\mathcal{M}J^{k-1}$, the induction hypothesis with $k - 1$ switches, and the definition (9) of I^k . To prove that we have equality between J^k and I^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 $\tau_0^* = t, \xi_0^* = i$ and

$$\begin{cases} \tau_\ell^* \triangleq \inf\{s \geq \tau_{\ell-1}^*: J^{k-\ell+1}(s, X_s^{t,x}, \xi_{\ell-1}^*) = \mathcal{M}J^{k-\ell}(s, X_s^{t,x}, i)\} \wedge T, \\ \xi_\ell^* \triangleq \arg \max_{j \neq i} (-C_{i,j} + J^{k-\ell}(\tau_\ell^*, X_{\tau_\ell^*}^{t,x}, j)), \quad \ell = 1, \dots, k. \end{cases} \quad (17) \quad \square$$

Observe that as an aside, the proof of Theorem 2 furnishes the *existence* of an optimal Markovian switching policy for I^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) in (17). 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 (9) 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 (2007), Bayraktar and Egami (2007). What is new, however, is its use for numerical computations as we do in Section 4.

Theorem 3. $\lim_{k \rightarrow \infty} J^k = J$ 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, $J^{k+1} = I^{k+1} \geq I^k = J^k$, and the monotone limit $J^\infty = \lim_{k \rightarrow \infty} J^k$, exists pointwise. Clearly, $J^\infty \leq J$ since $\mathcal{U}^k \subset \mathcal{U}$, so that $J \geq I^k = J^k$ for every integer $k \geq 1$. It remains to show that $J^\infty \geq J$. Fix $\epsilon > 0$, and let us assume that $u^\epsilon = ((\xi_1^*, \tau_1^*), \dots)$ is an ϵ -optimal policy for $J(t, x, i)$ in the sense that $J(t, x, i; u^\epsilon) \geq J(t, x, i) - \epsilon$. For any integer $k \geq 1$ we then define a policy $u^{k, \epsilon} \in \mathcal{U}^k(t)$ by $u_t^{k, \epsilon} = u_{t \wedge \tau_k^*}^\epsilon$. In other words, $u^{k, \epsilon}$ coincides with u^ϵ up to the k -th switch, after which it remains constant. Now

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

By admissibility of u^ϵ , as $k \rightarrow \infty$, both $\int_{\tau_{k+1}^*}^T \max_i |\psi_i(s, X_s^{t, x})| ds \rightarrow 0$ and $\sum_{\ell > k} C_{u_{\tau_\ell^*}^\epsilon, u_{\tau_\ell^*}^\epsilon} \rightarrow 0$. On the other hand,

$$\begin{aligned} \int_{\tau_{k+1}^*}^T \max_i |\psi_i(s, X_s^{t, x})| ds &\leq \int_{\tau_{k+1}^*}^T \max_i (1 + \psi_i(s, X_s^{t, x})^2) ds \\ &\leq \int_0^T \max_i (1 + \psi_i(s, X_s^{t, x})^2) ds. \end{aligned}$$

By Assumption 2, the latter expression is integrable and therefore by the dominated convergence theorem, the first expectation term in (18) tends to zero. Also, since $C_{i, j}$'s are strictly positive, by the monotone convergence theorem the second term in (18) also vanishes and we find

$$\lim_{k \rightarrow \infty} J(t, x, i; u^\epsilon) - J(t, x, i; u^{k, \epsilon}) = 0.$$

Using the fact that $u^{k, \epsilon} \in \mathcal{U}^k(t)$, so that $J^\infty(t, x, i) \geq J(t, x, i; u^{k, \epsilon})$ for any k , we obtain that $J^\infty(t, x, i) \geq J(t, x, i; u^\epsilon)$ which concludes the proof since ϵ was arbitrary. \square

Lemma 1. *The value function satisfies the coupled optimal stopping problem*

$$J(t, x, i) = \sup_{\tau \in \mathcal{S}_t} \mathbb{E} \left[\int_t^\tau \psi_i(s, X_s^{t, x}) ds + \mathcal{M}J(\tau, X_\tau^{t, x}, i) \right]. \quad (19)$$

In particular, the formal Dynamic Programming Principle holds for the value function, i.e. for any $t' \geq t$,

$$J(t, x, i) = \sup_{u \in \mathcal{U}(t, t')} \mathbb{E} \left[H([t, t'], x; u) + J(t', X_{t'}^{t, x}, u_{t'}) \mid X_t^{t, x} = x, u_t = i \right]. \quad (20)$$

Proof. By Theorem 3, $J(t, x, i) = \lim_{k \rightarrow \infty} \uparrow J^k(t, x, i)$. Since the right-hand-side of (19) is increasing in the terminal payoff $\mathcal{M}J$, we obtain

$$J^k(t, x, i) \leq \sup_{\tau \in \mathcal{S}_t} \mathbb{E} \left[\int_t^\tau \psi_i(s, X_s^{t, x}) ds + \mathcal{M}J(\tau, X_\tau^{t, x}, i) \right],$$

and taking the limit $k \rightarrow \infty$ on the right-hand-side one half of the result follows.

Conversely, let τ^* be an ϵ -optimal time for the right hand side of (19). By the monotone convergence theorem, we obtain

$$\begin{aligned} J(t, x, i) &= \lim_{k \rightarrow \infty} J^k(t, x, i) \\ &\geq \lim_{k \rightarrow \infty} \mathbb{E} \left[\int_t^{\tau^*} \psi_i(s, X_s^{t,x}) ds + \mathcal{M}J^{k-1}(\tau^*, X_{\tau^*}^{t,x}, i) \right] \\ &= \mathbb{E} \left[\int_t^{\tau^*} \psi_i(s, X_s^{t,x}) ds + \mathcal{M}J(\tau^*, X_{\tau^*}^{t,x}, i) \right] \\ &\geq \sup_{\tau \in \mathcal{S}_t} \mathbb{E} \left[\int_t^{\tau} \psi_i(s, X_s^{t,x}) ds + \mathcal{M}J(\tau, X_{\tau}^{t,x}, i) \right] - \epsilon. \end{aligned}$$

The dynamic programming principle follows by the strong Markov property of (X_t) and the Markovianity of ϵ -optimal stopping controls for (19) (Lepeltier and Marchal 1984). \square

From the proof of Theorem 3 it follows that for any $\epsilon > 0$ we can find an ϵ -optimal Markov strategy for $J(t, x, i)$ by looking at an optimal control of J^K for K large enough. The existence of (Markov) optimal policy for J is a much more delicate problem. General optimal stopping theory requires continuity of J to show existence of optimal control for (19). While we have regularity results for J^k , it is unknown whether $J^k \rightarrow J$ uniformly, which requires further assumptions on (X_t) .

As an example of how the latter can be achieved, we now directly show that for reasonable models the value function is continuous in x .

Lemma 2. *Suppose that (X_t) is continuous and the flow of (X_t) is Lipschitz, namely that μ and σ in (7) are continuously differentiable with bounded derivatives. Then the value functions $J(t, x, i)$ are locally Lipschitz continuous in x .*

Proof. The assumptions of the lemma imply that for any $p > 1$, there is a C_X such that $\mathbb{E}^x[\sup_t \|X_t\|^p] \leq C_X(1 + \|x\|^p)$ for x in a compact interval (Protter 2004, p. 342). The rest follows by a simple estimate and Gronwall's inequality. For x, y in a neighborhood of x_0 we find

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

for a constant $C = C(x_0)$. \square

3.3. Continuity of the Switching Boundary.

Lemma 3. *Suppose that the payoff rates ψ_m 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 horizon 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 $t < t'$,

$$\begin{aligned} |J(t, x, i) - J(t', x, i)| &= |J(0, x, i; T - t) - J(0, x, i; T - t')| \\ &\leq \sup_{u \in \mathcal{U}(T-t')} \mathbb{E} \left[\int_{T-t'}^{T-t} |\psi_{u_s}(X_s^{0,x})| ds \right] \\ &\leq (t' - t) \mathbb{E} \left[\sup_{0 \leq s \leq T} \max_i |\psi_i(X_s^{0,x})| \right], \end{aligned} \quad (21)$$

and the right hand side is uniform in x restricted to a bounded set by Assumption 2. \square

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

$$C_{i,j} = C_{i,k} + C_{k,j}, \quad (22)$$

it is also optimal to switch, at the same time τ , to regime j . Indeed, optimality of the switch at time τ from mode i to mode j implies that

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

and because of the additivity in (22) we obtain $-C_{k,j} + J(\tau, x, j) > J(\tau, x, k)$ which shows that switching to mode j is better than remaining in mode k at time τ . 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 $i > j$,

$$\frac{\partial(\psi_i - \psi_j)}{\partial x^n} \geq 0 \quad (\text{respectively } \leq 0). \quad (23)$$

Then it can be shown that $\partial_{x^n}(J(t, x, i) - J(t, x, j)) \geq 0$ (resp. ≤ 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

$$Switch(i, j; t) \triangleq \left\{ x \in \mathbb{R}^d : -C_{i,j} + J(t, x, j) > \max_{j' \neq j} \{-C_{i,j'} + J(t, x, j')\} \right\} \quad (24)$$

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 $Switch(i, j; t)$ as a function of t for different pairs (i, j) . By Lemma 3, the value function J is continuous in t , and consequently so is the switching set and its boundary. Condition (23) on the payoff rates is quite reasonable in practice, for instance all three of our examples in Section 5 satisfy it.

Remark 2. As shown above, an optimal strategy is described via hitting times by (X_t) of switching boundaries. Because of strictly positive switching costs, the switching boundaries should be separated from each other. If one can show that they are separated uniformly in x , then using the properties of jump-diffusions in (7), it is possible to derive locally uniform lower bounds on the τ_k^* and hence $J^k \rightarrow J$ uniformly on compacts. We leave this alternative method for studying the regularity of J to future work.

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 switching regions where the operational flexibility is immediately exercised. Let \mathcal{L}_X be the infinitesimal generator of (X_t) . Then roughly speaking (see Brekke and Øksendal (1994) for full details), one needs to study the *quasi-variational inequality*

$$\max \left(\mathcal{L}_X \phi(t, x, i) + \psi_i(t, x), -\phi(t, x, i) + \mathcal{M} \phi(t, x, i) \right) = 0, \quad (25)$$

subject to the terminal condition $\phi(T, x, i) = 0$. In other words, the function ϕ must dominate the barrier and satisfy the Kolmogorov partial differential equation (PDE) inside the continuation region. The so-called verification theorem states that a sufficiently smooth solution $\phi(t, x, i)$ of (25) 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 (25) (Pham 2007). 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 free boundaries. See for example Pham (2007) and the references therein for a discussion of smooth pasting. In the literature this assumption is often made based on heuristic arguments and even then 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. In contrast our probabilistic formulation sidesteps all these difficulties since we do not need 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 $X_0^{0,x} = x$ and let $Y_t^{k,i} \triangleq J^k(t, X_t^{0,x}, i)$, $\mathcal{M}_t^{k,i} = \mathcal{M}J^{k-1}(t, X_t^{0,x}, i)$. Note that $Y_t^{k,i}, \mathcal{M}_t^{k,i}$ are \mathcal{F}_t^X -measurable random quantities. Then there exists a triple $(Y^{k,i}, Z^{k,i}, A^{k,i})$ that solves

$$\begin{cases} Y_t^{k,i} = \int_t^T \psi_i(s, X_s^{0,x}) ds + A_T^{k,i} - A_t^{k,i} - \int_t^T Z_s^{k,i} \cdot dW_s, \\ Y_t^{k,i} \geq \mathcal{M}_t^{k,i}, \quad \text{with} \quad \int_0^T (Y_t^{k,i} - \mathcal{M}_t^{k,i}) dA_t^{k,i} = 0, \quad \text{and} \quad A_0^{k,i} = 0. \end{cases} \quad (26)$$

The first auxiliary process $Z^{k,i}$ satisfies $\mathbb{E}[\int_0^T \|Z_t^{k,i}\|^2 dt] < \infty$ and can be interpreted as a 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 was checked after (13) is necessary for (26) to make sense.

The elegance of (26) can be seen from the result of El-Karoui et al. (1997) who assert that subject to the above integrability conditions, there is a *unique* triple $(Y^{k,i}, Z^{k,i}, A^{k,i})$ solving (26) and necessarily satisfying $Y_0^{k,i} = J^k(0, x, i)$. In fact, denoting $(Y_s^{k,i}(t, x))_{t \leq s \leq T}$ the solution of (26) started with $X_t^{t,x} = x$, there exists a deterministic and jointly continuous function $u^{k,i}(t, x)$ such that $u^{k,i}(t, x) = Y_t^{k,i}(t, x)$. Moreover, this $u^{k,i}$ is known (El-Karoui et al. 1997) to be the unique viscosity solution of the QVI (25).

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 (26) is also continuous. Furthermore, analogues of stochastic flow theory imply that under regularity conditions on the forward SDE (7) giving the dynamics of (X_t) , $Y^{k,i}(t, x)$ will be continuously differentiable with respect to the initial condition x , and other parameters of (7). 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 3. Based on (19), the original value function $J(t, x, i)$ can be formally represented as a coupled system of reflected BSDE’s for $(Y^i)_{i \in \mathbb{Z}_M}$, $Y_t^i \equiv J(t, X_t^{0,x}, i)$:

$$\begin{cases} Y_t^i = \int_t^T \psi_i(s, X_s^{0,x}) ds + A_T^i - A_t^i - \int_t^T Z_s^i \cdot dW_s, \\ Y_t^i \geq \max_{j \neq i} \{-C_{i,j} + Y_t^j\}, \quad \text{and} \quad \int_0^T \left(Y_t^i - \max_{j \neq i} \{-C_{i,j} + Y_t^j\} \right) dA_t^i = 0. \end{cases} \quad (27)$$

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 (2007) explicitly proved existence of solutions to (27) 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. $\tau_\ell \in \mathcal{S}^\Delta$ and the new restricted set of admissible strategies is labelled \mathcal{U}^Δ . 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

$$J^k(t_1, X_{t_1}, i) = \max \left\{ \mathbb{E} \left[\int_{t_1}^{t_2} \psi_i(s, X_s) ds + J^k(t_2, X_{t_2}, i) \mid \mathcal{F}_{t_1}^X \right], \mathcal{M} J^{k-1}(t_1, X_{t_1}, i) \right\} \\ \simeq \left(\psi_i(t_1, X_{t_1}) \Delta t + \mathbb{E} [J^k(t_2, X_{t_2}, i) \mid \mathcal{F}_{t_1}^X] \right) \vee \left(\max_{j \neq i} \{ -C_{i,j} + J^{k-1}(t_1, X_{t_1}, j) \} \right). \quad (28)$$

We see that to find J^k it suffices to have a computationally efficient algorithm for evaluating the *conditional expectations* appearing in (28).

Recall that for a test function f , the conditional expectation $\mathbb{E}[f(X_{t_2}) \mid \mathcal{F}_{t_1}^X]$ is defined to be the $\mathcal{F}_{t_1}^X$ -measurable random variable F minimizing $\mathbb{E}[|f(X_{t_2}) - F|^2]$. On the other hand, since (X_t) is Markov, any such 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). Other alternatives are reviewed and compared in Section 5.3.

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 $L^2(\mathbb{P})$:

$$E_t(x) \simeq \hat{E}_t(x) = \sum_{j=1}^{N_B} \alpha_j B_j(x), \quad (29)$$

where $B_j(x)$ are the N_B bases and α_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 α_j after which we can use (29) to compute the conditional expectations.

Our contribution is to extend this approach to simultaneously compute all the J^k 's. 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 $(x_{m\Delta t}^n)$, $m = 0, \dots, M^\sharp$ 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 (28), for a given step $t_1 = m\Delta t$, $t_2 = t_1 + \Delta t$ and regime i we regress the known future

values $J^k(t_2, x_{t_2}^n, i)$ onto the current underlying values $\{B_j(x_{t_1}^n)\}$. As a result of regression we obtain a prediction $\hat{E}_{t_1}[J^k(t_1, \cdot, i)](x_{t_1}^n)$ for the continuation value along the n -th path. Comparing this with the current value $J^{k-1}(t_1, x_{t_1}^n, j)$ 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 $J^{k-1}(t_1, \cdot, \cdot)$ is known when computing $J^k(t_1, \cdot, \cdot)$. 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 advantageous to 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

$$J^k(m\Delta t, x, i) = \mathbb{E} \left[\sum_{j=m}^{\tau^k(m\Delta t, x, i)} \psi_i(j\Delta t, X_{j\Delta t}^{m\Delta t, x}) \cdot \Delta t + \mathcal{M}J^{k-1}(\tau^k\Delta t, X_{\tau^k\Delta t}^{m\Delta t, x}, i) \right]. \quad (30)$$

Then we have the analogue of (28) for τ^k :

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

With a slight abuse of notation, denote by $H^k(m\Delta t, x_{m\Delta t}^n, i)$ the realized pathwise gains, i.e. the actual cashflows of the operator along the simulated paths $(x_{m\Delta t}^n)$. Then combining (30) and (31), the recursive construction for H^k reads $H^k(T, x_T^n, i) = 0$,

$$H^k(m\Delta t, x_{m\Delta t}^n, i) = \begin{cases} \psi_i(m\Delta t, x_{m\Delta t}^n) \Delta t + H^k((m+1)\Delta t, x_{(m+1)\Delta t}^n, i), & \text{no switch;} \\ -C_{i,j} + H^{k-1}(m\Delta t, x_{m\Delta t}^n, j), & \text{switch to } j. \end{cases} \quad (32)$$

The value function J^k corresponds to average (expected value) of H^k , $J^k(t, x, i) = \hat{E}_t[H^k(t, \cdot, i)](x_t^n)$, and at $t = 0$ one recovers the value function in the same way as before, $J^k(0, x_0, i) = \frac{1}{N_p} \sum_n H^k(0, x_0^n, i)$. Finally, the set of paths on which we switch is given by $\{n: \hat{j}^n(m\Delta t; i) \neq i\}$ with

$$\hat{j}^n(m\Delta t; i) = \arg \max_j \left(-C_{i,j} + H^{k-1}(m\Delta t, x_{m\Delta t}^n, j), \right. \\ \left. \psi_i(m\Delta t, x_{m\Delta t}^n) \Delta t + \hat{E}_{m\Delta t}[H^k((m+1)\Delta t, \cdot, i)](x_{m\Delta t}^n) \right). \quad (33)$$

Observe that in this version the regression is used solely to update the optimal stopping times τ^k and the regressed values are never stored directly. This 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 H^k , and therefore also J^k .

The above description of the algorithm made direct use of the restricted value functions J^k . This makes the analysis of the algorithm more transparent and is also 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 using (19). Namely, we simply replace all J^k/H^k 's in (28)-(33) with just J/H . Hence, we regress $H(t_2, x_{t_2}^n, i)$ against $\{B_j(x_{t_1}^n)\}$ and use the resulting map \hat{E}_{t_1} to replace (33) with

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

4.2.1. *Choosing the Basis Functions.* The choice of appropriate basis functions (B_j) is 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 bases in Haugh and Kogan (2004). However, there is 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 ψ_i are usually linear and our favorite set of bases $B_j(x)$ are of the families $\{x^p\}$, $\{e^{\alpha x}\}$ 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 state (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 the examples below, N_B as small as 5 or 6 normally suffices, and having more bases often leads 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 δ . 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 δ . 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 $J(m\Delta t, \cdot, i)$ anyway; the advantage is that in discrete time, especially if δ is a multiple of Δt , 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 $\Delta t, M^\sharp, N_p, \bar{K}, \delta$. Set $D = \delta/\Delta t$ to be the time separation parameter.
- (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 pathwise gains and switching times $H^k(T, x_T^n, i) = 0, \tau^k(T, x_T^n, i) = M^\sharp, \forall i, k$.
- (4) Moving backward in time with $m = M^\sharp, \dots, 1$ repeat the Loop:
 - (a) Compute inductively the layers $k = 0, 1, \dots, \bar{K}$ using (11). For a mode i , compute the conditional expectation $\mathbb{E}[H^k(m\Delta t + \Delta t, \cdot, i) | \mathcal{F}_{m\Delta t}]$ by regressing $\{H^k(m\Delta t + \Delta t, x_{m\Delta t + \Delta t}^n, i)\}$ against current set of selected basis functions $\{B_j(x_{m\Delta t}^n)\}_{j=1}^{N_B}$ and adding the current reward $\psi_i(m\Delta t, x_{m\Delta t}^n) \cdot \Delta t$.
 - (b) Compare the obtained continuation value against the regressed switching values

$$\hat{E}_{m\Delta t}[H^{k-1}((m+D)\Delta t, \cdot, j)](x_{m\Delta t}^n) + \sum_{\ell=m}^{m+D} (\psi_j(\ell\Delta t, x_{\ell\Delta t}^n \cdot \Delta t) - C_{i,j}).$$

- (c) Select the optimal current mode based on (33).
- (d) Update the switching times and pathwise gains using (31), and (32).
- (5) end Loop.
- (6) Return $J^k(0, x_0^n, i) = \frac{1}{N_p} \sum_{n=1}^{N_p} H^k(\Delta t, x_{\Delta t}^n, i) + \psi_i(0, x_0) \cdot \Delta t$.
- (7) 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 (4d) directly uses the conditional expectations,

$$J^k(m\Delta t, x_{m\Delta t}^n, i) = \mathcal{M}J^{k-1}(m\Delta t, x_{m\Delta t}^n, i) \vee \left[\hat{E}_{m\Delta t} [J^k((m+1)\Delta t, \cdot, i)](x_{m\Delta t}^n) + \psi_i(m\Delta t, x_{m\Delta t}^n) \cdot \Delta t \right]. \quad (34)$$

4.3.1. Algorithm Requirements. The speed of the algorithm is $\mathcal{O}(M^2 \cdot \bar{K} \cdot N_p \cdot M^\sharp)$ where M is the number of possible scheduling modes, \bar{K} is maximum number of switches, N_p is the number of Monte Carlo paths, and M^\sharp is the number of discrete time steps used. Observe that during the main loop we only need to store the buffer $H(t, \cdot, \cdot), H(t + \Delta t, \cdot, \cdot), \dots, H(t + \delta, \cdot, \cdot)$ and $\tau(t, \cdot), \dots, \tau(t + \delta, \cdot)$. Consequently, the memory requirements are $\mathcal{O}(N_p \cdot (M^\sharp + (\frac{\delta}{\Delta t} + 1) \cdot \bar{K} \cdot M))$ where 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 includes four sources of errors: error due to discretizing the SDE (7), error due to restricting switching times to \mathcal{S}^Δ , 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 Δt 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 (34) 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 (26) and constructs a discrete time approximation to the triples $(Y^{k,i}, Z^{k,i}, A^{k,i})$, $k = 1, 2, \dots$. Recall that $Y_t^{k,i} = J^k(t, X_t^{t,x}, i)$, while $Z^{k,i}$ will be replaced with a conditional expectation similar to (28). To simplify the proofs we shall impose,

Assumption 3. *The functions $x \rightarrow \mu(x)$ and $x \rightarrow \sigma(x)$ appearing in (7) are $\mathcal{C}^2(\mathbb{R}^d)$ with uniformly bounded first derivatives. The payoff functions $\psi_m(t, x)$ are continuously differentiable with uniformly bounded derivatives.*

The convergence analysis will proceed in several steps, isolating each source of error identified above. Recall that time steps are indexed by $m = 0, \dots, M^\sharp$ and the paths by $n = 1, \dots, N \equiv N_p$. 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 $\|X\|_p = \mathbb{E}[|X|^p]^{1/p}$, $p \geq 1$. At this point let us also recall the following discrete version of Gronwall's inequality:

Lemma 4. *If positive sequences $(a_m), (b_m), (c_m)$, $0 \leq m \leq M^\#$ satisfy $a_{m-1} \leq \{(1 + C\Delta t)a_m + b_{m-1}\} \vee c_{m-1}$ then*

$$a_m \leq e^{C(T-m\Delta t)} \left[a_{M^\#} + \sum_{j \geq m} b_j + \max_{j \geq m} c_j \right]. \quad (35)$$

This lemma will be key for L^2 error 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 (7) with

$$X_t^\Delta = X_{t_1}^\Delta + \mu(X_{t_1}^\Delta)(t - t_1) + \sigma(X_{t_1}^\Delta) \cdot (W_t - W_{t_1}) + \sum_{i=N(t_1)}^{N(t)} \iota(X_{t_1}^\Delta, \xi_i), \quad \text{for } t \in (t_1, t_2]. \quad (36)$$

Note that for some models (X_t) can be simulated exactly; otherwise a standard L^p -bound for the discretization error is (Kloeden and Platen 1992)

$$\limsup_{\Delta t \rightarrow 0} \frac{1}{\sqrt{\Delta t}} \left\| \sup_{0 \leq t \leq T} |X_t - X_t^\Delta| \right\|_p < \infty \quad \forall p \geq 1. \quad (37)$$

Using (X_t^Δ) , the discretized versions labelled $(Y^{k,i,\Delta}, Z^{k,i,\Delta}, \mathcal{M}^{k,i,\Delta})$ of (26) are given by

$$\begin{cases} Y_{t_1}^{k,i,\Delta} = \mathcal{M}_{t_1}^{k,i,\Delta} \vee \left\{ \mathbb{E}[Y_{t_2}^{k,i,\Delta} | \mathcal{F}_{t_1}] + \psi_i(t_1, X_{t_1}^\Delta) \Delta t \right\}, & Y_T^{k,i,\Delta} = 0, \\ Z_{t_1}^{k,i,\Delta} = \mathbb{E} \left[Y_{t_2}^{k,i,\Delta} \cdot \frac{(W_{t_2} - W_{t_1})}{\Delta t} \middle| \mathcal{F}_{t_1} \right], \\ \mathcal{M}_{t_1}^{k,i,\Delta} = \max_{j \neq i} \left\{ -C_{i,j} + Y_{t_1}^{k-1,j,\Delta} \right\}. \end{cases} \quad (38)$$

The alternative

$$Y_{t_1}^{k,i,\Delta} = \max \left\{ \mathcal{M}_{t_1}^{k,i,\Delta}, \mathbb{E} \left[Y_{t_2}^{k,i,\Delta} + \psi_i(t_2, X_{t_2}^\Delta) \Delta t \middle| \mathcal{F}_{t_1} \right] \right\} \quad (39)$$

has also been used by some authors, e.g. Bally and Pagès (2003). In the numerical analysis terminology (38) is an explicit scheme, while (39) is an implicit scheme. From a financial point of view (38) means that the payoff between today and tomorrow is certain. It is not clear which assumption is more realistic but numerically for small Δt the differences seem to be negligible.

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

Theorem 4. *For all $p \geq 1$,*

$$\limsup_{\Delta t \rightarrow 0} \frac{1}{\sqrt{\Delta t}} \sup_{0 \leq m \leq M^\#} \left\| Y_{m\Delta t}^{k,i,\Delta} - Y_{m\Delta t}^{k,i} \right\|_p < \infty. \quad (40)$$

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 $\mathcal{O}(\Delta t)$, but the optimal stopping rules converge only with rate $\mathcal{O}(\sqrt{\Delta t})$, as in (40).

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 (38) by a finite-dimensional projection \mathcal{P}_{t_1} onto the orthonormal 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 $X_{t_1}^{0,x}$. Equation (38) becomes

$$\hat{Y}_{t_1}^{k,i} = \hat{\mathcal{M}}_{t_1}^{k,i} \vee \left(\mathcal{P}_{t_1}[\hat{Y}_{t_2}^{k,i}] + \psi_i(t_1, X_{t_1}^{\Delta}) \Delta t \right) \quad (41)$$

and the rest of (38) is changed accordingly.

Theorem 5. *Define $\hat{Y}_{m\Delta t}^{k,i}$ by recursively applying (41). Then*

- (1) $(\hat{Y}_{m\Delta t}^{k,i})$ is uniformly L^2 -integrable in m .
- (2) There exists a constant C such that

$$\left| \mathcal{P}_{t_1}[\hat{Y}_{t_2}^{k,i}] \right| \leq |\alpha_{t_1}| \cdot |B_{t_1}(X_{t_1}^{0,x})| \leq |B_{t_1}(X_{t_1}^{0,x})| \sqrt{C(1 + \|x\|^2)}. \quad (42)$$

The bound (42) allows us to give an *a priori* bound on the regression coefficients α_{t_1} . Namely, we can construct a truncation function T_{t_1} such that $|\mathcal{P}_{t_1}[\hat{Y}_{t_2}^{k,i}]| \leq T_{t_1}(X_{t_1})$ and $\mathbb{E}[T_{t_1}(X_{t_1})^2] < \infty$. This truncation implies that

$$|\alpha_{t_1}|^2 = \mathbb{E} \left| \mathcal{P}_{t_1}[\hat{Y}_{t_2}^{k,i}] \right|^2 \leq \mathbb{E} \left[T_{t_1}(X_{t_1})^2 \right].$$

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 $L^2(\mathcal{F}_{t_1})$. 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 $\mathcal{R}_{m\Delta t}(\mathbb{E}_{m\Delta t}[Y_{(m+1)\Delta t}^{k,i}])$ 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 6. *With the notation of (41) and $Y^{k,i} \equiv Y^{k,i,\Delta}$, there exists a constant C such that*

$$\begin{aligned} \max_{0 \leq m \leq M^\#} \mathbb{E} \left[|\hat{Y}_{m\Delta t}^{k,i} - Y_{m\Delta t}^{k,i}|^2 \right] &\leq C \sum_{m=0}^{M^\#} \mathbb{E} \left[\mathcal{R}_{m\Delta t}(\mathbb{E}_{m\Delta t}[Y_{(m+1)\Delta t}^{k,i}])^2 \right] \\ &\quad + \frac{C}{\Delta t} \max_{0 \leq m \leq M^\#} \max_{j \neq i} \mathbb{E} \left[|\hat{Y}_{m\Delta t}^{k-1,j} - Y_{m\Delta t}^{k-1,j}|^2 \right]. \end{aligned} \quad (43)$$

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 6.

Proposition 7. *There exists a constant C such that for any m, k, i ,*

$$\left\| \hat{Y}_{m\Delta t}^{k,i} - Y_{m\Delta t}^{k,i} \right\|_p \leq \frac{C}{\Delta t} \max_{m \leq j \leq M^\#} \left(\left\| (\mathbb{E}_{j\Delta t} - \mathcal{P}_{j\Delta t})(\hat{Y}_{(j+1)\Delta t}^{k,i}) \right\|_p + \max_{j' \neq i} \left\| \hat{Y}_{j\Delta t}^{k-1,j'} - Y_{j\Delta t}^{k-1,j'} \right\|_p \right). \quad (44)$$

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 $\mathcal{P}_{m\Delta t}$ 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

$$\hat{Y}_{t_1}^{k,i,n} = \hat{\mathcal{M}}_{t_1}^{k,i,n} \vee \left(T_{t_1}^n(\alpha_{t_1}^N \cdot B_{t_1}^n) + \psi_i(t_1, x_{t_1}^n) \Delta t \right), \quad (45)$$

$$\alpha_{t_1}^N = \arg \min_{\alpha} \frac{1}{N} \sum_{n=1}^N \left(\hat{Y}_{t_2}^{k,i,n} - \alpha \cdot B_{t_1}^n \right)^2. \quad (46)$$

The truncation $T_{t_1}^n(x)$ ensures that $|\alpha_{t_1}^N \cdot B_{t_1}^n| \leq 2 \cdot |\mathcal{P}_{t_1}[\hat{Y}_{t_2}^{k,i}]|$, which combined with (42) gives a priori control on the empirical solutions $\hat{Y}^{k,i,n}$. Note again that the formula (45) represents the TvR scheme (34) rather than the LS scheme (32).

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 (45) 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 α_{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 (45), 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. Even in simpler settings, the existing results in this direction are incomplete. For the LS scheme it has been shown that the use of optimal-stopping-time iteration in (31) 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 Δt is $\mathcal{O}(\log N \cdot N^{-1})$. However, the joint rate of convergence between the time step Δt , 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 $\mathcal{O}(\Delta t^{-1-d/4p} \cdot N^{-2p})$. However, the involved computational level is an order of magnitude higher than for our scheme that repeatedly re-uses the same paths. Alternatively, Gobet et al. (2005) obtain the $\mathcal{O}(N^{-1/2})$ L^2 -convergence rate when using a single set of paths but no reflection in (45). We conjecture that the L^2 sampling error is on the order $\mathcal{O}(\Delta t^{-1} N^{-1/2})$ 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

$$\begin{aligned} \|Y_{m\Delta t}^{k,i} - \hat{Y}_{m\Delta t}^{k,i,\Delta,N}\|_2 &\leq \|Y_{m\Delta t}^{k,i} - Y_{m\Delta t}^{k,i,\Delta}\|_2 + \|Y_{m\Delta t}^{k,i,\Delta} - \hat{Y}_{m\Delta t}^{k,i,\Delta}\|_2 + \|\hat{Y}_{m\Delta t}^{k,i,\Delta} - \hat{Y}_{m\Delta t}^{k,i,\Delta,N}\|_2 \\ &\leq C \left(\sqrt{\Delta t} + \frac{(\sum_{\ell=1}^k \sum_{j=m}^{M^\#-1} \mathbb{E}[\mathcal{R}_{j\Delta t}(\mathbb{E}_{j\Delta t}[Y_{(j+1)\Delta t}^{\ell,i}]^2)]^{1/2})}{(\Delta t)^k} + \frac{1}{\Delta t \sqrt{N}} \right). \end{aligned} \quad (47)$$

Even without the last term which is an unproven conjecture, these convergence rates are conservative. We observe a general explosion in the regression errors as we cascade through the number of switches k , cf. the second term in (47). 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) show that in the worst case one needs $\mathcal{O}(e^{N_B})$ Monte Carlo paths to obtain a comparable sampling error. It follows that to compute J^K we must use at least $N \sim (\Delta t)^{-K}$ 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 (45)), the pathwise realized gain 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 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. (A.3) and proof of (A.2).

From our experience, the algorithm is quite stable for up to a thousand time steps, $M^\# < 1000$. The corresponding discretization step Δt 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 as stated by (Egloff 2005), “the low but positive probability of ‘bad’ events dramatically worsens the worst-case-provable convergence rates”, 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 – 24000 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. We also see that in this example the value function was decreasing in N . 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 obscured by the projection and Monte Carlo errors whose bias is unknown.

TABLE 1. Convergence of Monte Carlo simulations for Example 1. True value was obtained by using a pde solver for (25).

# 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
<i>True</i>	5.93	–

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 $100 \cdot 8760 = 0.876$ terawatt. 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

$$\begin{cases} dP_t = P_t \left\{ 5(50 - \log P_t) dt + 0.5 dW_t^1 + d\left(\sum_{i=1}^{N(t)} (e^{\xi_i} - 1)\right) \right\} \\ dG_t = G_t \{2(6 - \log G_t) dt + 0.4(0.8 dW_t^1 + 0.6 dW_t^2)\} \end{cases} \quad (48)$$

where $\xi_i \sim \text{Exp}(10)$ and jump intensity is $\lambda = 8$ 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 $C_{ij} = 0.01G_t$.

We implement the LS algorithm of Section 4.3 using $N_p = 32,000$ paths, $M^\# = 360$ periods and $N_B = 6$ basis functions. A basic Matlab code without any optimization takes about five minutes to run on a 2.4GHz desktop and, while not lightning-fast, is acceptable for the back-office applications we envision. Furthermore, a practitioner is likely to have access to a lot more computing power, and one feature of Monte Carlo algorithms is that they are naturally parallelizable. We find

$$J(0, x_0, i) = [10.16, 10.22, 10.16], \quad i = 0, 1, 2.$$

These values indicate that initially the optimal regime is mode 1, regardless of the initial mode of the plant. To better understand the solution structure, Figure 1 presents two slices of the switching regions for different time instances t for example 1. We directly show the simulated pairs $(p_{m\Delta t}^n, x_{m\Delta t}^n)$, color-coded to indicate the corresponding optimal $\hat{j}^n(m\Delta t; i)$ in (33). As expected, the switching sets form three connected components in the (P_t, G_t) -space, but visualizing the respective switching boundaries for such a multidimensional (X_t) and several regimes is difficult. Also recall that the switching sets depend on remaining time to maturity $T - t$ due to the fixed switching costs and zero terminal condition that discourage mode changes close to T .

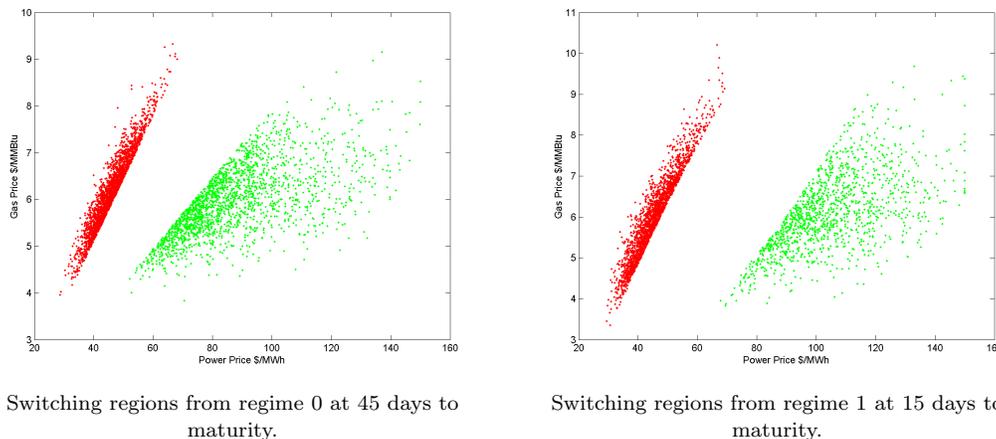


FIGURE 1. Empirical optimal switching regions for Example 1. Dark points represent the values $x_{m\Delta t}^n$ where regime 0 is optimal and light points represent $x_{m\Delta t}^n$ where regime 2 is optimal. The empty region in the middle represents the set where regime 1 is best. $N_p = 32,000$.

5.1. Approximating the Switching Boundary. The algorithm of Section 4.3 produces a natural approximation to the switching boundary. Recall that for each layer k and mode i , the switching boundary is given by the graph $\{(t, x_t)\}$, such that the minimal optimal switching time of $J^k(\cdot, \cdot, i)$ 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 switching time τ in (31), 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

$$\{x_{m\Delta t}^n : n \text{ is such that } \tau^k(m, x_{m\Delta t}^n, i) = m\} \quad (49)$$

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 (33). To illustrate this, the following example uses a simple one-dimensional (X_t) (that can be interpreted as the spark spread), where switching boundaries can be plotted directly as a function of time.

Example 2. 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

$$\{\psi_0(X_t) = 0, \quad \psi_1(X_t) = 10(X_t - 10)\},$$

and the switching cost between them is $C_{0,1} = C_{1,0} = 0.3$.

Figure 2 illustrates the results for Example 2. The top panel of Figure 2 shows a realized path of (X_t) superimposed over the switching boundaries of $J(t, x, i)$. 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 the *at-the-money* price is $x = 10$). Thus, if the plant is ‘off’, it will be brought online only when X_t is about 10.8. When $10 \leq X_t \leq 10.8$, and the mode is ‘off’, it is myopically better to run the plant, but the fixed switching costs combine with future uncertainty to give a dynamic preference to continue in regime 0. 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 make effect of time insignificant.

Using the top panel of Figure 2, we construct an optimal strategy u^* as follows. Given an initial regime u_0 and a realization of $(X_t^{0,x})$, we keep track for each $j \neq u_0$ of the switching sets $Switch(u_0, j; t)$ defined in (24). At the first instant t that $X_t^{0,x} \in Switch(u_0, j; t)$ for some j , we invoke our operational flexibility, switch to regime j and proceed inductively until T . For example in Figure 2 there is a total of four such mode switches, starting from regime $u_0 = 1$. The bottom panel of Figure 2 shows the total cumulative wealth obtained from this policy. The discrete switching costs are indicated by drops of $C = 0.3$ at the switching times.

5.2. Three-Dimensional Example.

Example 3. 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 liquified natural gas infrastructure in US coastal 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 (48) by keeping same

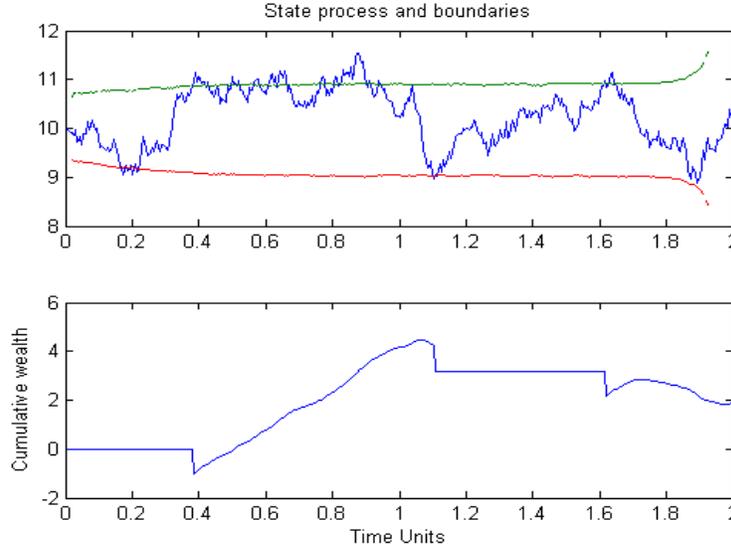


FIGURE 2. Implementing optimal policy for the one-dimensional example 2. We start with $X_0 = 10, u_0 = 1$.

equations for (P_t) and (G_t) and adding

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

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 over-capacity, which produces more power but is less efficient (due to excessive heat loss, higher wear-and-tear, etc.). Thus, we assume that there are a total of five regimes available to the operator: complete shutdown $\psi_0(X_t) \equiv 0$, as well as (all values are in millions of dollars and using annualized units):

$$\left\{ \begin{array}{ll} \psi_1(X_t) = P_t - 8G_t, & \psi_2(X_t) = P_t - O_t, \\ \psi_3(X_t) = 3P_t - 32G_t, & \psi_4(X_t) = 3P_t - 4O_t. \end{array} \right\}$$

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:

- (1) Number of Regimes. The more operational regimes are available to the agent, 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.
- (2) Switching Delay. Longer switching delays will lower the tolling value since the agent faces more risk when changing the operational regime.
- (3) Number of Switches. Many plants have *cycling* restrictions due to engineering constraints. As a result, the agent 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 by comparing the above scenarios to the base case. 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%. This illustrates the need for dynamic optimization, as regimes 3 and 4 would be completely excluded from conventional NPV calculation. Restrictions on number of switches is a typical dynamic “volumetric” constraint that is very difficult to incorporate in a strip-of-options method. We find that allowing only two switches $k = 2$, results in a 20% loss of value. For $k = 7$, one loses only about 1%; put differently, in the given model, one expects to make less than seven switches over the next six months in nearly every case. Finally, the switching delay causes losses on the order of 2-4%, which might be due to 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 3. We ran the LS scheme with $M^\# = 360$ steps, $\Delta t = 1/2$ day, $N_p = 24000$ paths and $\delta = 2$ days. Initial prices are $P_0 = 50, G_0 = 6, O_0 = 50$.

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

5.3. Alternative Computational Methods. The optimal switching problem can be solved using a variety of other numerical approaches. First, one can directly tackle the quasi-variational formulation of (25). A number of 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 a discrete-space approximation (\tilde{X}_t) . One can then apply the dynamic programming approach directly, since conditional expectations are transformed into weighted averages using the transition probabilities of (\tilde{X}_t) . This approach includes the (b) Markov Chain approximation method of Kushner and Dupuis (2001) and the (c) Optimal Quantization technique of Bally et al. (2005). Finally, one can go back to the key representation (28) 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 (25) was the most accurate but also the slowest approach. Let us stress that the benchmarking was carried out on a relatively simple one-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) (a practical implementation is likely to use several factors to model the evolution of power prices, another several factors to model gas prices, etc.). 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. Instead, we recommend the use of the LS and quantization schemes. 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.

Remark 4. 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. The implementation for our heap of optimal stopping problems is beyond the scope of this paper.

5.4. 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” $\{m\Delta t\}$, this means that the overall value of running the power plant is approximated by

$$V(0, x; \Delta t) = \mathbb{E} \left[\sum_{m=0}^{M^\#} (P_{m\Delta t} - \overline{HR} \cdot G_{m\Delta t} - K)^+ \cdot \Delta t \mid (P_0, G_0) = (p_0, g_0) \right]. \quad (51)$$

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

Lemma 5. *Let $\bar{C} = \max_{i,j} C_{i,j}$. In the limit of vanishing switching costs,*

$$\lim_{\bar{C} \rightarrow 0} J(0, x, i) = V(0, x).$$

Proof. Fix a time step Δt 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,

$$V(t_1, x) = \mathbb{E} \left[V(t_2, X_{t_2}^{t_1, x}) \right] + (P_{t_1} - \overline{HR} \cdot G_{t_1} - K)^+ \cdot \Delta t. \quad (52)$$

Since $\psi_0(x) = 0$, and $\psi_1(x) = (P_{t_1} - \overline{HR} \cdot G_{t_1} - K)$, for any x , there is a $i(x)$ such that $\psi_{i(x)}(x) = (P_{t_1} - \overline{HR} \cdot G_{t_1} - K)^+$. For this mode $i(x)$, $J(t_1, x, i(x)) \geq \mathbb{E} [J(t_2, X_{t_2}^{t_1, x}, i(x))] + \psi_{i(x)}(x) \Delta t$, and combining with (52),

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

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

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

Inducting on t_1 , we get $\sup_x \sup_t |J(t, x, i) - V(t, x)| \leq M^\# \cdot \bar{C}$ and taking the limit $\bar{C} \rightarrow 0$ the lemma follows. \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 as hard as the original problem.

6. EXTENSIONS.

6.1. 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 $\tilde{\lambda}$. Letting \tilde{T}_k represent the random time of the k -th emergency, we have $\mathbb{P}[\tilde{T}_k \geq s + \tilde{T}_{k-1}] = e^{-\tilde{\lambda}s}$. Since times between outages have the memoryless exponential distribution, $\tilde{\lambda}$ acts like an additional discount factor and the recursive construction (11) becomes

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

The first term above represents the probability of an outage and a forced switch to mode ‘off’ before τ . 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 $\mathbb{E}[(1 - \tilde{\lambda}\Delta t)J^k(t + \Delta t, X_{t+\Delta t}, i) + \tilde{\lambda}\Delta t(J^{k-1}(t + \Delta t, X_{t+\Delta t}, 0) - C_{i,0}) | \mathcal{F}_t] + \psi_i(t, X_t)\Delta t$.

6.2. 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

$$\begin{cases} J(t, x, c, i) = \sup_{u \in \mathcal{U}(t)} \mathbb{E} \left[\int_t^T \psi_{u_s}(s, X_s^{t,x}) ds - \sum_{\tau_k < T} C_{u_{\tau_k}, u_{\tau_k}} \middle| I_t = c, u_t = i \right] \\ J(t, x, 0, i) \equiv 0, \quad I_t \geq 0. \end{cases} \quad (53)$$

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

The new complexity in (53) 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 τ denotes the time of next switching decision then (53) can be re-written as

$$J(t, x, c, i) = \sup_{\tau \in \mathcal{S}_t, j \neq i} \mathbb{E} \left[\int_t^\tau \psi_i(s, X_s^{t,x}) ds + J(\tau, X_\tau^{t,x}, I_\tau, j) - C_{i,j} \middle| I_t = c \right], \quad (54)$$

and $I_\tau = c - a_i(\tau - t)$, allowing us to effectively eliminate (I_t) . As a result, given Δt 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 4. 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 copper 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

$$\{\psi_0(X_t) = -0.5e^{-0.04t}, \quad \psi_1(t, X_t) = e^{-0.04t}(10(X_t - 0.5) - 5(X_t - 0.5)^+)\}.$$

The switching cost is $C_{0,1} = C_{1,0} = 0.2$; 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 $J(m\Delta t, x, i) \geq 0$ at each time point.

The original paper of (Brennan and Schwartz 1985, p. 343) solved this 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, taking 600 time steps, 10000 paths and 75 grid points 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.

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 '07	1.35	4.41	8.15	12.44	17.52	22.41

TABLE 3. Comparison of pricing methods for the mining example of Brennan and Schwartz (1985), varying the initial copper price P_0 . The initial inventory is $I_0 = 150$ million pounds and the values are in millions of dollars.

6.3. Price Impact. As mentioned in the introduction, the price impact of operator's actions is another important feature. In the simplest model, the agent moves the market by a deterministic amount. Such an effect can be directly incorporated into the actual profit functions ψ_m . A more sophisticated effect can be obtained by relying on supply-demand equilibrium. Assume that the demand D_t is stochastic and given in terms of exogenous state variables, such as temperature. The supply consists of a fixed exogenous supply stack $S(P_t)$, which determines the amount of power supplied given the electricity price P_t , as well as an endogenous part ξ_t controlled by the operator. The overall clearing price is then determined by solving $S(P_t) + \xi_t = D_t$ or $P_t := S^{-1}(D_t - \xi_t)$. Choosing $X_t = (D_t, G_t)$ we have $\psi_m(X_t) = S^{-1}(D_t - \xi_m) - HR \cdot G_t$, 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 in return for a fixed payment. This leads to volumetric risk with respect to (D_t) .

It is also possible to imagine that actions of the agent influence the *parameters* of the price processes, e.g. the mean-reversion level of P_t in (8). Such a model can also be dealt with straightforwardly. Notice that we just need to simulate (X_t) under each regime $i \in \mathbb{Z}_M$ and use the LS scheme over each set of paths $(x_{m\Delta t}^{n,i})$, $i \in \mathbb{Z}_M$. Because $x_{m\Delta t}^{n,i} \neq x_{m\Delta t}^{n,j}$, we must replace (32) with

$$H^k(m\Delta t, x_{m\Delta t}^{n,i}, i) = \begin{cases} \psi_i(m\Delta t, x_{m\Delta t}^{n,i}) \Delta t + H^k((m+1)\Delta t, x_{(m+1)\Delta t}^{n,i}, i), & \text{no switch;} \\ -C_{i,j} + \hat{E}[H^{k-1}(m\Delta t, \cdot, j)](x_{m\Delta t}^{n,i}), & \text{switch to } j. \end{cases}$$

where the second case resets the pathwise values $\{H^{k-1}(m\Delta t, x_{m\Delta t}^{n,j}, j)\}$ based on the j -regime paths. This effectively blends the LS scheme for the continuation values and the TvR scheme for the switching values; such a mixed method for American options has also been described by Egloff (2005).

7. 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.

From the point of view of realistic implementation, it is clear that the models presented so far are simplified. 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.

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 2007). 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 useful to have further analysis of the structure of the switching sets, extending Section 3.3.

ACKNOWLEDGMENT

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

APPENDIX: PROOFS FOR SECTION 4.4.

Proof of Theorem 4. We proceed by induction on k . Let $\mathcal{S}^\Delta(m\Delta t) = \{\tau \in \mathcal{S} : \tau \in \{j\Delta t, m \leq j \leq M^\#\}\}$ 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 3 (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

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

or $\|R_{m\Delta t}^{k,i} - Y_{m\Delta t}^{k,i,\Delta}\|_p \leq C\sqrt{\Delta t} + \|\max_{0 \leq j \leq M^\#} \max_{j' \neq i} |Y_{j\Delta t}^{k-1,j'} - Y_{j\Delta t}^{k-1,j',\Delta}|\|_p$ by (37). 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 ψ_m 's. By induction on k , the $\mathcal{O}(\sqrt{\Delta t})$ error propagates through for any fixed layer k . \square

Proof of Theorem 5. We first check that for a fixed k , $\hat{Y}_{m\Delta t}^{k,i}$ is uniformly L^2 -integrable in time. Using

$$\mathbb{E}|(X_1 \vee Y_1) - (X_2 \vee Y_2)|^2 \leq (1 + \gamma)\mathbb{E}|X_1 - X_2|^2 \vee \left(1 + \frac{1}{\gamma}\right)\mathbb{E}|Y_1 - Y_2|^2 \quad (\text{A.1})$$

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

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

so that by Gronwall's lemma

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

Completing the induction on k , $\max_{0 \leq m \leq M^\#} \mathbb{E}\left[|\hat{Y}_{m\Delta t}^{k,i}|^2\right] \leq C(\Delta t)^{-k}(1 + \|x\|^2)$. This provides a priori estimates on the regression result $\mathcal{P}_{t_1}[\hat{Y}_{t_2}^{k,i}]$. By (A.2) and orthonormality of (B_{t_1}) we have

$$\mathbb{E}\left[\mathcal{P}_{t_1}[\hat{Y}_{t_2}^{k,i}]^2\right] = \mathbb{E}\left[|\alpha_{t_1}|^2 |B_{t_1}|^2\right] = |\alpha_{t_1}|^2 \leq C(1 + \|x\|^2)$$

and (42) follows. \square

Proof of Theorem 6. The theorem shows that the regression errors add up when moving across a fixed layer k and get multiplied by $\frac{1}{\Delta t}$ when moving down to layer $k-1$. 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

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

Combined once again with Lemma 4 this is enough to prove the theorem. To show the claimed inequality, re-write

$$\begin{aligned} \eta_{t_1}^{k,i} &= \mathbb{E}\left[\left\{\mathcal{P}_{t_1}(\hat{Y}_{t_2}^{k,i}) + \psi_i(t_1, X_{t_1})\Delta t \vee \hat{\mathcal{M}}_{t_1}^{k,i}\right\} - \left\{\mathbb{E}_{t_1}(Y_{t_2}^{k,i}) + \psi_i(t_1, X_{t_1})\Delta t \vee \mathcal{M}_{t_1}^{k,i}\right\}\right]^2 \\ &\leq \mathbb{E}\left[|\mathcal{P}_{t_1}(\hat{Y}_{t_2}^{k,i}) - \mathbb{E}_{t_1}(Y_{t_2}^{k,i})|^2 \vee |\hat{\mathcal{M}}_{t_1}^{k,i} - \mathcal{M}_{t_1}^{k,i}|^2\right] \end{aligned} \quad (\text{A.3})$$

Since \mathcal{P}_{t_1} is an L^2 -projection, $\mathcal{P}_{t_1}(Z) = \mathcal{P}_{t_1}(\mathbb{E}_{t_1}[Z])$ for any 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

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

by orthogonality of the remainder \mathcal{R}_{t_1} . Substituting into (A.3) and using (A.1) we get

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

which immediately implies (43). \square

REFERENCES

- L. Andersen and M. Broadie. A primal-dual simulation algorithm for pricing multi-dimensional American options. *Management Sci.*, 50(9):1222–1234, 2004.
- V. Bally and G. Pagès. Error analysis of the optimal quantization algorithm for obstacle problems. *Stochastic Process. Appl.*, 106(1):1–40, 2003. ISSN 0304-4149.
- V. Bally, G. Pagès, and J. Printems. A quantization tree method for pricing and hedging multidimensional American options. *Math. Finance*, 15(1):119–168, 2005.
- A. Bar-Ilan, A. Sulem, and A. Zanello. Time-to-build and capacity choice. *J. Econom. Dynam. Control*, 26(1):69–98, 2002. ISSN 0165-1889.
- E. Bayraktar and M. Egami. On the one-dimensional optimal switching problem, 2007. working paper.
- A. Bensoussan and J.-L. Lions. *Impulse Control and Quasi-Variational Inequalities*. Gauthier-Villars, Paris, 1984.
- B. Bouchard and N. Touzi. Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Process. Appl.*, 111(2):175–206, 2004. ISSN 0304-4149.
- K. Brekke and B. Øksendal. Optimal switching in an economic activity under uncertainty. *SIAM J. Control Optim.*, 32(4):1021–1036, 1994. ISSN 0363-0129.
- M. Brennan and E. Schwartz. Evaluating natural resource investments. *J. Business*, 58:135–157, 1985.
- R. Carmona and V. Durrleman. Pricing and hedging spread options. *SIAM Rev.*, 45(4):627–685, 2003. ISSN 1095-7200.
- R. Carmona and M. Ludkovski. Valuation of energy storage: An optimal switching approach. Technical report, University of Michigan, 2007. Submitted.
- R. Carmona and N. Touzi. Optimal multiple stopping and valuation of swing options, 2005.
- J. F. Carrière. Valuation of the early-exercise price for options using simulations and nonparametric regression. *Insurance : Math. Econom.*, 19:19–30, 1996.
- D. Chevance. Numerical methods for backward stochastic differential equations. In L. Rogers and D. Talay, editors, *Numerical methods in finance*, pages 232–244. Cambridge University Press, 1997.
- E. Clément, D. Lamberton, and P. Protter. An analysis of a least squares regression algorithm for American option pricing. *Finance Stoch.*, 6:449–471, 2002.
- S. Deng and Z. Xia. Pricing and hedging electric supply contracts: a case with tolling agreements, 2005. preprint.
- A. Dixit. Entry and exit decisions under uncertainty. *J. Polit. Economy*, 97(3):620–638, June 1989.
- P. Dupuis and H. Wang. On the convergence from discrete to continuous time in an optimal stopping problem. *Ann. Appl. Probab.*, 15:1339–1366, 2005.
- D. Egloff. Monte Carlo algorithms for optimal stopping and statistical learning. *Ann. Appl. Probab.*, 15(2):1396–1432, 2005.
- N. El-Karoui, C. Kapoudjian, E. Pardoux, S. Peng, and M. Quenez. Reflected solutions of backward SDE's and related obstacle problems for PDE's. *Ann. Probab.*, 25:702–737, 1997.
- A. Eydeland and K. Wolyniec. *Energy and Power Risk Management: New Developments in Modeling, Pricing and Hedging*. John Wiley&Sons, Hoboken, NJ, 2003.

- P. Glasserman and B. Yu. Number of paths versus number of basis functions in American option pricing. *Ann. Appl. Probab.*, 14(4):2090–2119, 2004.
- E. Gobet, J. Lemor, and X. Warin. A regression-based Monte-Carlo method to solve backward stochastic differential equations. *Ann. Appl. Probab.*, 15(3):2172–2002, 2005.
- X. Guo and H. Pham. Optimal partially reversible investment with entry decision and general production function. *Stochastic Process. Appl.*, 115:705–736, 2005.
- S. Hamadène and M. Jeanblanc. On the starting and stopping problem: application in reversible investments. *Mathematics of Operations Research*, 32(1):182–192, 2007.
- S. Hamadène and J. Lepeltier. Reflected BSDEs and mixed game problem. *Stochastic Process. Appl.*, 85(2):177–188, 2000. ISSN 0304-4149.
- M. Haugh and L. Kogan. Pricing American options: A duality approach. *Oper. Res.*, 52(2):258–270, March-April 2004.
- M. Jeanblanc and A. Shiryaev. Optimization of the flow of dividends (in russian). *Russian Math Surveys*, 50: 257–277, 1995.
- P. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*, volume 23 of *Applications of Mathematics*. Springer-Verlag, New York, 1992.
- H. Kushner and P. Dupuis. *Numerical methods for stochastic control problems in continuous time*, volume 24 of *Applications of Mathematics*. Springer-Verlag, New York, second edition, 2001.
- J. Lepeltier and B. Marchal. Théorie générale du contrôle impulsif markovien. *SIAM J. Control Optim.*, 22 (4):645–665, 1984. ISSN 0363-0129.
- F. Longstaff and E. Schwartz. Valuing American options by simulations: a simple least squares approach. *Review Finan. Stud.*, 14:113–148, 2001.
- M. Ludkovski. Financial hedging of operational flexibility, 2007. Working Paper, University of Michigan.
- M. Ludkovski. *Optimal Switching with Application to Energy Tolling Agreements*. PhD thesis, Princeton University, 2005.
- N. Meinshausen and B. Hambly. Monte Carlo methods for the valuation of multiple exercise options. *Math. Finance*, 14:557–583, 2004.
- B. Øksendal and A. Sulem. *Applied stochastic control of jump diffusions*. Springer-Verlag, Berlin, 2005.
- H. Pham. On the smooth-fit property for one dimensional optimal switching problem. In *Séminaire de Probabilités, vol XL*. 2007.
- H. Pham and V. Ly Vath. Explicit solution to an optimal switching problem in the two-regime case. *SIAM J. Control Optim.*, 46(2):395–426, 2007.
- P. Protter. *Stochastic integration and differential equations*, volume 21 of *Applications of Mathematics*. Springer-Verlag, Berlin, 2004. ISBN 3-540-50996-8. 2nd Edition.
- J. N. Tsitsiklis and B. Van Roy. Regression methods for pricing complex American-style options. *IEEE Trans. Neural Netw.*, 12(4):694–703, July 2001.
- H. Wang. A sequential entry problem with forced exits. *Math. Oper. Res.*, 30:501–520, 2005.
- P. Wilmott, S. Howison, and J. Dewynne. *The mathematics of financial derivatives*. Cambridge University Press, Cambridge, 1995. ISBN 0-521-49699-3; 0-521-49789-2. A student introduction.
- A. Yushkevich. Optimal switching problem for countable Markov chains: average reward criterion. *Math. Methods Oper. Res.*, 53(1):1–24, 2001. ISSN 1432-2994.
- M. Zervos. A problem of sequential entry and exit decisions combined with discretionary stopping. *SIAM J. Control Optim.*, 42(2):397–421, 2003. ISSN 1095-7138.

(R. Carmona) DEPARTMENT OF OPERATIONS RESEARCH AND FINANCIAL ENGINEERING, BENDHEIM CENTER FOR FINANCE, PRINCETON UNIVERSITY, PRINCETON, NJ 08544 USA
E-mail address: rcarmona@princeton.edu

(M. Ludkovski) DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MICHIGAN, ANN ARBOR, MI 48109, USA
E-mail address: mludkov@umich.edu, www.umich.edu/mludkov