International Journal of Pure and Applied Mathematics
Volume 87 No. 4 2013, 559-586
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
doi: http://dx.doi.org/10.12732/ijpam.v87i4.7
ijpam.eu

OPTIMAL STOPPING FOR PERPETUAL AMERICAN FORWARDS UNDER MARKOVIAN SWITCHING REGIME

Thomas Seaquist[1], Andrzej Korzeniowski[2 §]
[1,2]Department of Mathematics
University of Texas at Arlington
P.O. Box 19408, Arlington, TX 76019, USA

Abstract: We derive and solve a system of ODEs over the infinite time horizon which maximizes the payout for a forward contract under a real world measure modulated by n-state continuous time Markov chain.

AMS Subject Classification: 90J27, 60J70, 91G20, 93E20
Key Words: Markovian switching, semimartingales, Ito's formula, logarithmic version of Black-Scholes' equations, optimal stopping

1. Introduction

In this paper we discuss the pricing of one the most typical claims on a stock, the forward contract, referred to from here on as a forward. A forward is a contract in which one party agrees to give another a share of stock for a fixed price at an agreed upon date in the future. The stock is being sold forward. We will price this contract under the framework of Markovian switching of average rate of return and volatility. Once this derivative is priced under a martingale measure, a new infinite time horizon American version of this forward will be discussed. This extends upon the work of [2] for a two state model. It will be shown that this new contract cannot be priced under a risk neutral measure.

Received: June 9, 2013

© 2013 Academic Publications, Ltd.
url: www.acadpubl.eu

[§]Correspondence author

Despite this, we will discuss the optimal exercise strategy and expected payoff of this contract under the actual real world probability measure whose expected return and volatility are usually estimated from historical data.

2. Pricing Forwards in a Markov Modulated Market

The pricing of a forward goes as follows: what price K should one party be obligated to pay for one share of stock $\{S_t\}_{t\geq 0}$ at a predetermined time T in the future in order to make the game fair for both parties? Mathematically, we say

$$E[e^{-rT}(S_T - K)] = 0$$

where S_0 is the present value of a share of stock. It is well known that in order to avoid arbitrage K must be priced under the risk neutral or martingale measure. This is the measure under which the discounted stock process is a martingale and is found by Girsanov change of measure. The stock process described by $dS_t = \mu S_t dt + \sigma S_t dW_t$ becomes $dS_t = rS_t dt + \sigma S_t dW_t$ under the risk neutral (martingale) measure. Thus we have

$$E[e^{-rT}(S_T - K)] = S_0 - Ke^{-rT}$$

and it becomes clear that the proper value is $K = S_0 e^{rT}$.

Now we will focus on how a Markov modulated market changes the dynamics of the stock process. We start with the stock process

$$dS_t = \mu(\xi_t)S_t dt + \sigma(\xi_t)S_t dW_t$$

where ξ_t is a regular Markov chain representing the n states of the market. After switching to the risk neutral measure we get the stock process

$$dS_t = r(\xi_t)S_t dt + \sigma(\xi_t)S_t dW_t.$$

As before, we find K by mandating that the discounted expected payoff be zero.

$$\begin{aligned} 0 &= E[e^{-\int_0^T r(\xi_s)ds}(S_t - K)] \\ &= S_0 - E[e^{-\int_0^T r(\xi_s)ds}]K \end{aligned}$$

giving

$$K = \frac{S_0}{E[e^{-\int_0^T r(\xi_s)ds}]}.$$

All that remains is to find the value of $E[e^{-\int_0^T r(\xi_s)ds}]$.

Proposition 2.1. *Define*

$$M(t,i) = E^i[e^{-\int_0^t r(\xi_s)ds}].$$

$M(t,i)$ *satisfies the following ODE system,*

$$\mathbf{M}' = (Q - R)\mathbf{M} \quad M(0,i) = 1$$

where $\mathbf{M}$ *is the vector whose* i*-th element is* $M(t,i)$*,* Q *is the infinitesimal generator of* ξ_t*, and* R *is the diagonal matrix whose* i*-th diagonal element is* $r(i)$*.*

Proof. We will derive the ODE for $M(t,i)$ by conditioning on the first transition time and a time Δt. Below T_1 is the time of the first transition and P_{ij} is the probability of ξ_t transition from i to j given a transition at t.

$$\begin{aligned}
M(t,i) &= E^i[e^{-\int_0^t r(\xi_s)ds}] \\
&= E^i[e^{-r(i)\Delta t} e^{-\int_{\Delta t}^t r(\xi_s)ds} \mathbb{1}_{\{T_1 > \Delta t\}}] + E^i[e^{-r(i)T_1} e^{-\int_{T_1}^t r(\xi_s)ds} \mathbb{1}_{\{T_1 \le \Delta t\}}] \\
&= E^i\left[E[e^{-r(i)\Delta t} e^{-\int_{\Delta t}^t r(\xi_s)ds} \mathbb{1}_{\{T_1 > \Delta t\}} \mid \mathcal{F}_{\Delta t}]\right] \qquad (1) \\
&\qquad + E^i\left[E[e^{-r(i)T_1} e^{-\int_{T_1}^t r(\xi_s)ds} \mathbb{1}_{\{T_1 \le \Delta t\}} \mid \mathcal{F}_{T_1}]\right] \\
&= \mathbb{1}_{\{T_1 > \Delta t\}} e^{-r(i)\Delta t} E^i\left[E[e^{-\int_{\Delta t}^t r(\xi_s)ds} \mid \mathcal{F}_{\Delta t}]\right] \qquad (2) \\
&\qquad + E^i\left[\mathbb{1}_{\{T_1 \le \Delta t\}} e^{-r(i)T_1} E[e^{-\int_{T_1}^t r(\xi_s)ds} \mid \mathcal{F}_{T_1}]\right] \\
&= \mathbb{1}_{\{T_1 > \Delta t\}} e^{-r(i)\Delta t} E^i\left[E[e^{-\int_{\Delta t}^t r(\xi_s)ds}]\right] + E^i\left[\mathbb{1}_{\{T_1 \le \Delta t\}} e^{-r(i)T_1} E[e^{-\int_{T_1}^t r(\xi_s)ds}]\right] \\
& \qquad (3)
\end{aligned}$$

$$\begin{aligned}
&= P(T_1 > \Delta t) e^{-r(i)\Delta t} M(t - \Delta t, i) + E^i\Big[\mathbb{1}_{\{T_1 \le \Delta t\}} e^{-r(i)T_1} \sum_{\substack{j=1 \\ j \ne i}}^n P_{ij} M(t - T_1, j)\Big] \\
&= e^{q_{ii}\Delta t} e^{-r(i)\Delta t} M(t - \Delta t, i) + P(T_1 \le \Delta t) e^{-r(i)\eta\Delta t} \sum_{\substack{j=1 \\ j \ne i}}^n P_{ij} M(t - \eta\Delta t, j) \qquad (4)
\end{aligned}$$

$$= e^{(-r(i)+q_{ii})\Delta t}M(t-\Delta t,i) + (1-e^{q_{ii}\Delta t})e^{-r(i)\eta\Delta t}\sum_{\substack{j=1\\ j\neq i}}^{n} P_{ij}M(t-\eta\Delta t,j)$$

where $\eta \in (0,1)$. Now (4) will be justified. Define

$$f(s) := e^{-r(i)s}\sum_{\substack{j=1\\ j\neq i}}^{n} P_{ij}M(t-s,j)$$

and notice that $f(s)$ is continuous since $M(s,j)$ is continuous. Since

$$E^i[\mathbb{1}_{\{T_1\leq\Delta t\}}f(T_1)] = \int_0^{\Delta t} f(s)dF_{T_1}(s)$$

where F_{T_1} is the distribution of T_1, we apply the mean value theorem for integrals to get

$$\begin{aligned} E^i[\mathbb{1}_{\{T_1\leq\Delta t\}}f(T_1)] &= f(\eta\Delta t)\Big(F_{T_1}(\Delta t) - F_{T_1}(0)\Big) \\ &= f(\eta\Delta t)P(T_1 \leq \Delta T) \end{aligned}$$

where $\eta \in (0,1)$, thus providing justification for the above.

Next we use a Taylor expansion around $t-\Delta t$ for $M(t,i)$ to arrive at

$$\begin{aligned} &M(t-\Delta t,i) + M'(t-\Delta t,i)\Delta t + o(\Delta t) \\ &\quad = e^{(-r(i)+q_{ii})\Delta t}M(t-\Delta t,i) + (1-e^{q_{ii}\Delta t})e^{-r(i)\eta\Delta t}\sum_{\substack{j=1\\ j\neq i}}^{n} P_{ij}M(t-\eta\Delta t,j). \end{aligned}$$

Subtracting $M(t,i)$ from both sides, dividing by $-\Delta t$, and reorganizing yields

$$\begin{aligned} &\frac{M(t-\Delta t,i)-M(t,i)}{-\Delta t} - M'(t-\Delta t,i) - \frac{o(\Delta t)}{\Delta t} \\ &= \frac{e^{(-r(i)+q_{ii})\Delta t}M(t-\Delta t,i)-M(t,i)}{-\Delta t} \\ &\qquad\qquad + \frac{(e^{q_{ii}\Delta t}-1)}{\Delta t}e^{-r(i)\eta\Delta t}\sum_{\substack{j=1\\ j\neq i}}^{n} P_{ij}M(t-\eta\Delta t,j) \\ &= \frac{M(t-\Delta t,i)-M(t,i)}{-\Delta t} - M(t-\Delta t,i)\frac{e^{(-r(i)+q_{ii})\Delta t}-1}{\Delta t} \end{aligned}$$

$$+\frac{(e^{q_{ii}\Delta t}-1)}{\Delta t}e^{-r(i)\eta\Delta t}\sum_{\substack{j=1\\ j\neq i}}^{n}P_{ij}M(t-\eta\Delta t, j).$$

Taking the limit $\Delta t \to 0$ we get the ODE

$$0 = M'(t,i) - (-r(i) + q_{ii})M(t,i) + \sum_{\substack{j=1\\ j\neq i}}^{n} q_{ii}P_{ij}M(t, j).$$

Using that $q_{ii}P_{ij} = -q_{ij}$,

$$\begin{aligned} M'(t,i) &= -r(i)M(t,i) + \left(\sum_{\substack{j=1\\ j\neq i}}^{n} q_{ij}M(t, j)\right) + q_{ii}M(t,i) \\ &= -r(i)M(t,i) + \sum_{j=1}^{n} q_{ij}M(t, j). \end{aligned}$$

This is written nicely in matrix form,

$$\mathbf{M}' = (Q - R)\mathbf{M}$$

where Q is the infinitesimal generator of ξ_t and R is the diagonal matrix whose i-th diagonal element is $r(i)$. Since

$$M(0,i) = E^i[e^{-\int_0^0 r(\xi_s)ds}] = 1,$$

the proof is complete. □

The solution of the ODE in proposition 2.1 is determined by the eigenvalues and eigenvectors of $Q - R$:

$$\mathbf{M}(t) = \sum_{i=1}^{n} C_i\mathbf{V_i}e^{\lambda_i t}$$

where V_i and λ_i are the i-th corresponding eigenvector and eigenvalue and the C_i's make up n arbitrary constants completely determined by the initial condition. For complex eigenvalues, we interpret the solution as the sum of sines and cosines in the standard way.

Now we can find the value K, dependent upon the initial Markov state and stock value:

$$K = \frac{S_0}{E[e^{-\int_0^T r(\xi_s)ds}]} = \frac{S_0}{M(T,\xi_0)}.$$

We have now priced a forward in a Markov modulated market.

3. Optimal Exercise Strategy for American Style Forwards

In this section we will develop a contract which will be referred to as a "perpetual American future". The contract is formed by one party agreeing to buy a share of stock for a price K at any time of his choosing in the future. Perpetual refers to the fact that the contract has an infinite time horizon and thus no fixed termination time. American meaning that the contract can be exercised at any time. The payoff of the contract at a time t is given by $(S_t - K)$.

Next we attempt to find the fair value of K. We must keep in mind one major difference: the contract can be exercised at any time. Ideally the contract will be exercised at an optimal time, *i.e.*, the time that maximizes expected future discounted profit. We conclude that the expected profit under the real world measure P from the contract is

$$\sup_\tau E_P[e^{-r\tau}(S_\tau - K)\mathbb{1}_{\tau<\infty}]$$

the indicator being necessary since if the contract is never exercised, we gain no profit. Without the indicator $\mathbb{1}_{\tau<\infty}$, it is unclear how to interpret $\lim_{t\to\infty} e^{-rt}(S_t - K)$. If we are to price this derivative on the open market, we must use the the risk neutral measure Q. We set

$$\sup_\tau E_Q[e^{-r\tau}(S_\tau - K)\mathbb{1}_{\tau<\infty}] = 0$$

to find K. It must be determined whether such a value exist.

We now have two questions to answer. The first, what is the optimal exercise strategy and expected profit under the real measure for the owner of a contract. The second, is there a fair price for K under the risk neutral measure. Both of these questions will be answered when an optimal stopping time is identified. Notice that the value of the contract is not dependent on time since there is no termination date. Thus we conclude that the optimal stopping time must be only dependent upon the stock price. We conclude that if an optimal stopping time exists, it is of the form

$$\tau_c = \inf\{t \geq 0 : S_t \geq c\}$$

for some threshold c. Lets first answer the question of expected profit under this optimal exercise strategy:

$$E_P[e^{-r\tau_c}(S_{\tau_c} - K)\mathbb{1}_{\tau_c<\infty}] = E_P[e^{-r\tau_c}\mathbb{1}_{\tau_c<\infty}](c - K).$$

We need only find the value of $E_P[e^{-r\tau_c}\mathbb{1}_{\tau_c<\infty}]$ to determine the expected profit.

First, we will take the perspective of the logarithmic stock process, which by Ito's formula we find to be

$$\begin{aligned} X_t &= \ln(S_t) \\ &= X_0 + (\mu - \frac{1}{2}\sigma^2)t + \sigma W_t. \end{aligned}$$

The hitting time for S_t hitting c is the same as X_t hitting $\ln(c)$. Define $b = \ln(c)$ and $\alpha = \mu - \frac{1}{2}\sigma^2$. Define the stopping time

$$T_b = \inf\{t \geq 0 : X_t \geq b\}$$

We have

$$\begin{aligned} E_P[e^{-r\tau_c}\mathbb{1}_{\tau_c<\infty}] &= E_P[e^{-rT_b}\mathbb{1}_{T_b<\infty}] \\ &= \int_0^\infty e^{-rt} f(t)dt \end{aligned}$$

where $f(t)$ is the density function of T_b. All that needs to be done is to find the density function. Notice that the last line above is the Laplace transform of the density function. This will allow us to find the expected profit of the contract. The following two results will provide the tools necessary to find the density function of T_b.

Lemma 3.1. *Let* $M_t = \max\{W_s : 0 \leq s \leq t\}$.

$$P(M_t \geq b) = 2P(W_t \geq b) = 2\left(1 - \Phi\left(\frac{b}{\sqrt{t}}\right)\right)$$

where Φ *is standard normal distribution.*

For a proof we refer the reader to [4] (Theorem 3.15, pg 71).

Theorem 3.1. *Let* $X_t = x + \alpha t + \sigma W_t$. *The distribution of the hitting time* $T_b = \inf\{t \geq 0 : X_t \geq b\}$ *is*

$$F(t) = 1 - \Phi\left(\frac{b - x - \alpha t}{\sigma\sqrt{t}}\right) - e^{\frac{2\alpha(b-x)}{\sigma^2}}\Phi\left(\frac{b - x - \alpha t}{\sigma\sqrt{t}}\right), \quad x < b, \quad t > 0$$

and the density function is given by

$$f(t) = \frac{(b-x)}{\sqrt{2\pi t^3 \sigma^2}} \exp\left(\frac{-(b-x-\alpha t)^2}{2t\sigma^2}\right), \quad x < b, \quad t > 0.$$

Before we proceed to the proof of this theorem, observe that if $\alpha > 0$, then we have a proper probability distribution, but if $\alpha < 0$ we have a defective probability distribution, i.e.,

$$\lim_{t\to\infty} F(t) = e^{\frac{2\alpha(b-x)}{\sigma^2}} < 1.$$

Thus there would be a positive probability the arithmetic Brownian motion never hits level b when $\alpha < 0$. It is now clear that is very important that we include $\mathbb{1}_{T_b<\infty}$ in taking expected values.

Proof. In the case that $\alpha = 0$ we use lemma 3.1 to show that

$$\begin{aligned} P(T_b \le t) &= P(\max\{X_s : 0 \le s \le t\} \ge b) \\ &= P\left(M_t \ge \frac{b-x}{\sigma}\right) \\ &= 2\left(1 - \Phi\left(\frac{b-x}{\sigma\sqrt{t}}\right)\right). \end{aligned} \tag{5}$$

Next Girsanov's theorem will be used to remove the drift from X_t in the case that α is not zero. Define the measure Q by a Radon-Nikodym derivative:

$$\frac{dQ}{dP} = \exp(-\frac{1}{2}\frac{\alpha^2}{\sigma^2}t - \frac{\alpha}{\sigma}W_t).$$

Under the new measure Q, $\tilde{W}_t = W_t + \frac{\alpha}{\sigma}t$ is a Wiener process. Now the distribution of T_b will be found. Define $G_t = \frac{dP}{dQ} = \exp(-\frac{1}{2}\frac{\alpha^2}{\sigma^2}t + \frac{\alpha}{\sigma}\tilde{W}_t)$. We use the fact that $E_P[H] = E_Q[HG_T]$ for any bounded Borel function H and where G_T is the terminal value of the martingale G_s.

$$\begin{aligned} P(T_b \le t) &= E_P[\mathbb{1}_{\{T_b \le t\}}] \\ &= E_Q[\mathbb{1}_{\{T_b \le t\}} G_T] \\ &= E_Q[E[\mathbb{1}_{\{T_b \le t\}} G_T \mid \mathcal{F}_{T_b \wedge t}]] \\ &= E_Q[\mathbb{1}_{\{T_b \le t\}} G_{T_b \wedge t}] \\ &= E_Q[\mathbb{1}_{\{T_b \le t\}} G_{T_b}] \end{aligned}$$

$$= E_Q[\mathbb{1}_{\{T_b \le t\}} \exp\left(-\frac{1}{2}\frac{\alpha^2}{\sigma^2}T_b + \frac{\alpha}{\sigma}(\frac{b-x}{\sigma})\right)]$$

$$= \int_0^t \exp(-\frac{1}{2}\frac{\alpha^2}{\sigma^2}s + \frac{\alpha}{\sigma}(\frac{b-x}{\sigma})f(s)ds \qquad (6)$$

where $f(s)$ is the density function of T_b under Q, *i.e.* with no drift. From (5), we get

$$f(s) = \frac{d}{ds}2\Phi\left(\frac{b-x}{\sigma\sqrt{s}}\right) = \frac{(b-x)}{\sqrt{2\pi t^3\sigma^2}}\exp\left(\frac{-(b-x)^2}{2t\sigma^2}\right).$$

Differentiating (6) and simplifying we get the density function

$$f(t) = \frac{(b-x)}{\sqrt{2\pi t^3\sigma^2}}\exp\left(\frac{-(b-x-\alpha t)^2}{2t\sigma^2}\right).$$

It is apparent upon differentiation that $F'(t) = f(t)$ and the proof is finished. □

All that remains is to find the Laplace transform of $f(s)$ and we have $E[e^{-rT_b}\mathbb{1}_{T_b<\infty}]$.

$$E[e^{-rT_b}\mathbb{1}_{T_b<\infty}] = \int_0^\infty e^{-rs}f(s)ds$$
$$= \exp(-\frac{b-x}{\sigma^2}(\sqrt{\alpha^2+2r\sigma^2}-\alpha)).$$

Now we will answer the question of whether there is a fair value for K to put this derivative on the market. In the risk neutral measure Q, $\alpha = r - \frac{1}{2}\sigma^2$ and we get

$$E_Q[e^{-rT_b}\mathbb{1}_{T_b<\infty}] = e^{-(b-x)}.$$

Lets find the value of b which maximizes profit,

$$E_Q[e^{-rT_b}(S_{T_b} - K)\mathbb{1}_{T_b<\infty}] = e^{-(b-x)}(e^b - K) = e^x(1 - Ke^{-b}).$$

To maximize profit $b \to \infty$, thus it is impossible to price this derivative on the market. However, it is still useful to address the issue of optimal exercise of the contract supposing that one has such a contract. Even if the contract can't be sold freely on the market, it is still feasible that the contract might be given to an individual. For example, an employer might give such a contract to an employee as a form of payment or bonus.

Lets now address the issue of optimal exercise and expected profit. For $\mu > r$

$$E^x[e^{-rt}(S_t - K)] = e^{(\mu-r)t}e^x - e^{-rt}K$$

in which case we want to hang on to this contract, because the discounted expected profit will gain value with time. Because of this, there is no optimal stopping time.

Next we address the case when $\mu < r$ when it is not in our best interest to keep the contract indefinitely. Lets find the optimal level for b

$$E^x[e^{-rT_b}(S_{T_b} - K)\mathbb{1}_{T_b<\infty}] = \exp\left(-\frac{b-x}{\sigma^2}(\sqrt{\alpha^2 + 2r\sigma^2} - \alpha)\right)(e^b - K)$$

for $x < b$. Maximizing this relative to b yields

$$e^b = \frac{K}{1 - \frac{\sigma^2}{z}}$$

where $z = \sqrt{\alpha^2 + 2r\sigma^2} - \alpha$ and $\alpha = \mu - \frac{1}{2}\sigma^2$. Notice that when $\mu < r$ then $\frac{\sigma^2}{z} < 1$ and there is a valid value for b. On the other hand if $\mu \geq r$ then $\frac{\sigma^2}{z} \geq 1$ and there does not exist a value for b as we would expect from the previous discussion.

Finally the expected discounted profit when $\mu < r$ is given by

$$\begin{aligned} E^x[e^{-rT_b}(S_{T_b} - K)\mathbb{1}_{T_b<\infty}] &= E^x[e^{-rT_b}\mathbb{1}_{T_b<\infty}]\left(\frac{K}{1 - \frac{\sigma^2}{z}} - K\right) \\ &= K\exp\left(-\frac{z(b-x)}{\sigma^2}\right)\left(\frac{\sigma^2}{z - \sigma^2}\right). \end{aligned}$$

We have identified an optimal stopping strategy along with an accompanying discounted expected profit, which concludes this section.

4. Generalized Ito's Formula Including a Pure Jump Markov Process

Here we develop the mathematical tools necessary for the later needed analysis when introducing Markov modulation into the American forward contract. A generalized Ito-formula will be presented and used to find the infinitesimal generator of the two dimensional process (X_t, ξ_t). These two results will then be combined to present a very elegant version of Ito's formula with a Markov chain.

The general Ito's formula for an n-tuple of possibly discontinuous semimartingales can be found in a monograph of stochastic integration ([6], p. 74, Th. 33), and is given below.

Theorem 4.1 (Generalized Ito's Formula [6]). *Let $X = (X^1, \dots, X^n)$ be collection of semimartingales and let $f : \mathbb{R}^n \to \mathbb{R}$ have continuous second order partial derivatives. Then $f(X_t)$ is a semimartingale and the following formula holds:*

$$f(X_t) = f(X_0) + \sum_{i=1}^{n} \int_{0^+}^{t} \frac{\partial f}{\partial x_i}(X_{s^-}) dX_s^i + \frac{1}{2} \sum_{1 \le i,j}^{i,j \le n} \int_{0^+}^{t} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_{s^-}) d[X^i, X^j]_s^c$$
$$+ \sum_{0<s\le t} \left\{ \Delta f(X_s) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(X_{s^-}) \Delta X_s^i \right\},$$

where, $X_{t^-} = \lim_{t\to 0^-} X_t$, $\Delta X_t = X_t - X_{t^-}$ *and* $\Delta f(X_t) = f(X_t) - f(X_{t^-})$. $[X,X]_t$ *and* $[X,Y]_t$ *denote the quadratic variation and the quadratic covariation respectively and* $[X,Y]_t^c$ *denotes the path by path continuous part of* $[X,Y]_t$.

Now Ito's Formula will be applied to the function $f(Z_t)$ where $Z_t = (X_t, \xi_t)$. Recall that X_t is the Markov modulated Ito-process defined

$$dX_t = \left(\mu(\xi_t) - \frac{1}{2}\sigma^2(\xi_t)\right) dt + \sigma(\xi_t) dW_t$$

and ξ_t is the n state Markov chain.

$$f(Z_t) = f(Z_0) + \int_{0^+}^{t} f_x(Z_{s^-}) dX_s + \int_{0^+}^{t} f_\xi(Z_{s^-}) d\xi_s + \frac{1}{2} \int_{0^+}^{t} f_{xx}(Z_{s^-}) d[X,X]_s^c$$
$$+ \int_{0^+}^{t} f_{x\xi}(Z_{s^-}) d[X,\xi]_s^c + \frac{1}{2} \int_{0^+}^{t} f_{\xi\xi}(Z_{s^-}) d[\xi,\xi]_s^c$$
$$+ \sum_{s\le t} [\Delta f(Z_s) - f_x(Z_{x^-}) \Delta X_s - f_\xi(Z_{s^-}) \Delta \xi_s].$$

Since ξ_s is of bounded variation and X_s is continuous, $[X,\xi]_s^c = [\xi,\xi]_s^c = \Delta X_s = 0$. In addition, since ξ_s is a pure jump process, $\int_{0^+}^{t} f_\xi(Z_{s^-}) d\xi_s = \sum_{s\le t} f_\xi(Z_{s^-}) \Delta \xi_s$. By $dX_s = (\mu(\xi_s) - \frac{1}{2}\sigma^2(\xi_s)) ds + \sigma(\xi_s) dW_s$ and $d[X,X]_s^c = \sigma^2(\xi_s) ds$ one obtains

$$f(Z_t) = f(Z_0) + \int_{0}^{t} \left[\left(\mu(\xi_s) - \frac{1}{2}\sigma^2(\xi_s)\right) f_x(Z_s) + \frac{1}{2}\sigma^2(\xi_s) f_{xx}(Z_s) \right] ds$$
$$+ \int_{0^+}^{t} \sigma(\xi_{s^-}) f_x(Z_{s^-}) dW_s + \sum_{s\le t} \Delta f(Z_s).$$

Remark 4.1. The left limit was ignored in the ds integral since $X_s \neq X_{s^-}$ at only a finite number of times almost surely. Left limits cannot be ignored in the Ito integral since the integrand without it would not be a predictable process.

To simplify notation and for clarity, the observation is made that the infinitesimal operator L_X of the process X_t is

$$L_X[f](x,i) = \left(\mu(i) - \frac{1}{2}\sigma^2(i)\right) f_x(x,i) + \frac{1}{2}\sigma^2(i) f_{xx}(x,i),$$

and yields

$$f(Z_t) = f(Z_0) + \int_0^t L_X[f](Z_s)ds + \int_{0^+}^t \sigma(\xi_{s^-}) f_x(Z_{s^-})dW_s + \sum_{s\le t} \Delta f(Z_s). \tag{7}$$

Proposition 4.1. *The infinitesimal generator of the process $Z_t = (X_t, \xi_t)$ for a bounded function $f(\cdot,\xi) \in C^2(\mathbb{R})$is given by*

$$L_{(X,\xi)}[f](x,i) = L_X[f](x,i) + L_\xi[f](x,i)$$

where

$$L_X[f](x,i) = \left(\mu(i) - \frac{1}{2}\sigma^2(i)\right) f_x(x,i) + \frac{1}{2}\sigma^2(i) f_{xx}(x,i)$$
$$L_\xi[f](x,i) = \sum_{j=1}^n q_{ij} f(x,j)$$

and q_{ij} is the infinitesimal transition rate from state i to j.

Proof. The generator is defined by

$$L_{(X,\xi)}[f](x,i) = \lim_{t\searrow 0} \frac{E^{(x,i)}[f(X_t,\xi_t)] - f(x,i)}{t}.$$

We apply Ito's Formula (7) and observe that by standard properties of Ito integrals $\int_{0^+}^t \sigma(\xi_{s^-}) f_x(Z_{s^-})dW_s$ is a martingale and thus $E[\int_{0^+}^t \sigma(\xi_{s^-}) f_x(Z_{s^-})dW_s] = 0$. Now

$$\begin{aligned}
L_{(X,\xi)}[f](x,i) &= \lim_{t\searrow 0} \frac{E^{(x,i)}[\int_0^t L_X[f](X_s,\xi_s)ds + \sum_{s\le t} \Delta f(X_s,\xi_s)]}{t} \\
&= L_X[f](x,i) + \lim_{t\searrow 0} \frac{E^{(x,i)}[\sum_{s\le t} \Delta f(X_s,\xi_s)]}{t} \\
&= L_X[f](x,i) + \lim_{t\searrow 0} \frac{E^{(x,i)}\left[(f(X_T,\xi_t) - f(X_T,i))\mathbb{1}_{\{N(t)\le 1\}}\right]}{t}
\end{aligned}$$

$$+\lim_{t\searrow 0}\frac{E^{(x,i)}\left[\sum_{s\le t}\Delta f(X_s,\xi_s)\mathbb{1}_{\{N(t)\ge 2\}}\right]}{t} \tag{8}$$

where $N(t)$ counts the number of jumps of ξ_t and T is the time of the first jump. Next, it will be shown that

$$\lim_{t\searrow 0}\frac{1}{t}E^{(x,i)}[\sum_{s\le t}\Delta f(X_t,\xi_t)\mathbb{1}_{\{N(t)\ge 2\}}]=0.$$

It was assumed that f is a bounded function, so let $\sup_{(x,i)}|f(x,i)|\le M$. Let $N^*(t)$ be a Poisson process with rate $\lambda=\max_i(-q_{ii})$, recalling that $-q_{ii}$ is the rate of leaving state i. It is clear to see that choosing this maximal rate yields the inequality $E[N(t)]\le E[N^*(t)]$. Now we have that

$$\begin{aligned}E^{(x,i)}\left[\left|\sum_{s\le t}\Delta f(X_s,\xi_s)\mathbb{1}_{\{N(t)\ge 2\}}\right|\right] &\le E[2MN(t)\mathbb{1}_{\{N(t)\ge 2\}}]\\ &\le E[2MN^*(t)\mathbb{1}_{\{N(t)\ge 2\}}]\\ &=2M\sum_{k=2}^{\infty}kP(N^*(t)=k)\\ &=2M\Big(E[N^*(t)]-P(N^*(t)=1)\Big)\\ &=2M\lambda t(1-e^{-\lambda t})\end{aligned}$$

which gives

$$\lim_{t\searrow 0}\frac{E^{(x,i)}[|\sum_{s\le t}\Delta f(X_s,\xi_s)\mathbb{1}_{\{N(t)\ge 2\}}|]}{t}\le\lim_{t\searrow 0}2M\lambda(1-e^{-\lambda t})=0$$

and thus

$$\lim_{t\searrow 0}\frac{E^{(x,i)}[\sum_{s\le t}\Delta f(X_s,\xi_s)\mathbb{1}_{\{N(t)\ge 2\}}]}{t}=0.$$

Applying the above in (8) we have

$$\begin{aligned}L_{(X,\xi)}[f](x,i) &= L_X[f](x,i)+\lim_{t\searrow 0}\frac{E^{(x,i)}\left[\Big(f(X_T,\xi_t)-f(X_T,i)\Big)\mathbb{1}_{\{N(t)\le 1\}}\right]}{t}\\ &= L_X[f](x,i)+\lim_{t\searrow 0}\frac{E^{(x,i)}\left[E\Big[\Big(f(X_T,\xi_t)-f(X_T,i)\Big)\mathbb{1}_{\{N(t)\le 1\}}\mid X_t\Big]\right]}{t}\end{aligned}$$

$$= L_X[f](x,i) + \lim_{t \searrow 0} E^{(x,i)} \left[\sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{P_{ij}(t)}{t} \Big(f(X_T, j) - f(X_T, i) \Big) \mathbb{1}_{\{N(t) \leq 1\}} \right]$$

$$\begin{aligned}
&= L_X[f](x,i) + E^{(x,i)} \left[\sum_{\substack{j=1 \\ j \neq i}}^{n} \lim_{t \searrow 0} \frac{P_{ij}(t)}{t} \Big(f(X_T, j) - f(X_T, i) \Big) \mathbb{1}_{\{N(t) \leq 1\}} \right] \\
&= L_X[f](x,i) + \sum_{\substack{j=1 \\ j \neq i}}^{n} q_{ij} \Big(f(x,j) - f(x,i) \Big) \\
&= L_X[f](x,i) + \sum_{\substack{j=1 \\ j \neq i}}^{n} q_{ij} f(x,j) - \sum_{j=1}^{n} q_{ij} f(x,i) \\
&= L_X[f](x,i) + \sum_{\substack{j=1 \\ j \neq i}}^{n} q_{ij} f(x,j) + q_{ii} f(x,i) \\
&= L_X[f](x,i) + \sum_{j=1}^{n} q_{ij} f(x,j) \\
&= L_X[f](x,i) + L_\xi[f](x,i)
\end{aligned}$$

where $P_{ij}(t) = P(\xi_t = j|\xi_0 = i)$, and q_{ij} is defined by $\lim_{t \to 0} \frac{P_{ij}(t)}{t}$. The interchange of limit and expectation is justified since f and $\frac{P_{ij}(t)}{t}$ are bounded and thus Lebesgue's dominated convergence theorem applies. □

From Proposition 1.7 ([1], p. 162) we have

$$f(X_t, \xi_t) - f(X_0, \xi_0) - \int_0^t L_{(X,\xi)}[f](X_s, \xi_s) ds$$

is a martingale with respect to the natural filtration generated by (X_t, ξ_t). Combining Proposition 4.1 and Ito's formula (7)

$$\int_{0^+}^t \sigma(\xi_{s^-}) f_x(Z_{s^-}) dW_s + \sum_{s \leq t} \Delta f(X_s, \xi_s) - \int_0^t L_\xi[f](X_s, \xi_s) ds$$

is a martingale. On the other hand, by standard properties of Ito-integrals, $\int_0^t \sigma(\xi_s) f_x(X_s, \xi_s) dW_s$ is a martingale, thus

$$M_t^f := \sum_{s\le t} \Delta f(X_s, \xi_s) - \int_0^t L_\xi[f](X_s, \xi_s) ds$$

is a martingale. This produces a nice semimartingale decomposition:

$$\sum_{s\le t} \Delta f(X_s, \xi_s) = \int_0^t L_\xi[f](X_s, \xi_s) ds + M_t^f.$$

As a result, Ito's formula provides a key representation for the process (X_t, ξ_t) as follows:

$$\begin{aligned} f(X_t, \xi_t) = f(X_0, \xi_0) + \int_0^t L_{(X,\xi)}[f](X_s, \xi_s) ds + \int_{0^+}^t \sigma(\xi_{s^-}) f_x(Z_{s^-}) dW_s \quad (9) \\ + M_t^f \quad (10) \\ = f(X_0, \xi_0) + \int_0^t L_{(X,\xi)}[f](X_s, \xi_s) ds + Martingale. \end{aligned}$$

Another useful result that will be needed later is to apply Ito's Formula to the function defined $F(e^{-\int_0^t r(\xi_s)ds}, Z_t) := e^{-\int_0^t r(\xi_s)ds} f(Z_t)$. Applying Theorem 4.1 in the previous manner, we obtain

$$\begin{aligned} & e^{-\int_0^t r(\xi_s)ds} f(Z_t) \quad (11) \\ & = f(Z_0) + \int_0^t f(Z_s) d\left(e^{-\int_0^s r(\xi_u)du}\right) + \int_0^t e^{-\int_0^s r(\xi_u)du} L_{(X,\xi)}[f](Z_s) ds \\ & \qquad + \int_{0^+}^t e^{-\int_0^s r(\xi_u)du} \sigma(\xi_{s^-}) f_x(Z_{s^-}) dW_s + M_t^F \\ & = f(Z_0) + \int_0^t e^{-\int_0^s r(\xi_u)du} \Big(L_{(X,\xi)}[f](Z_s) - r(\xi_s) f(Z_s)\Big) ds \\ & \qquad + \int_{0^+}^t e^{-\int_0^s r(\xi_u)du} \sigma(\xi_{s^-}) f_x(Z_{s^-}) dW_s + M_t^F \quad (12) \\ & = f(Z_0) + \int_0^t e^{-\int_0^s r(\xi_u)du} \Big(L_{(X,\xi)}[f](Z_s) - r(\xi_s) f(Z_s)\Big) ds + Martingale. \end{aligned}$$

5. American Style Forwards in a Markov Modulated Market

In this section we will allow the average return and volatility of the stock process change between n states by a Markov chain. However, here we fix the risk-free interest rate r. Our goal is to identify an optimal stopping strategy for the perpetual American style forward discussed in Section 3 and to find the discounted expected payoff under the optimal strategy.

Recall that the Markov modulated stock process is defined

$$dS_t = \mu(\xi_t)S_t dt + \sigma(\xi_t)S_t dW_t$$

as before. Here we do not apply a change of measure since we are interested in the real world optimal strategy and payoff. We define the optimal payoff as

$$V(x,i) = \sup_{\tau} E^{(s,i)}[e^{-r\tau}(S_\tau - K)\mathbb{1}_{\{\tau<\infty\}}]$$

again applying the indicator to make it clear that there is zero payoff if the contract is held indefinitely. We have

$$S_t = S_0 \exp\left(\int_0^t \left(\mu(\xi_s) - \frac{1}{2}\sigma^2(\xi_s)\right) ds + \int_0^t \sigma(\xi_s) dW_s\right).$$

and

$$S_0 \exp\left(\int_0^t -\frac{1}{2}\sigma^2(\xi_s) ds + \int_0^t \sigma(\xi_s) dW_s\right)$$

is a martingale.

Proposition 5.1. *Let* $X_t = \ln(S_t)$*, then*

$$dX_t = \left(\mu(\xi_t) - \frac{1}{2}\sigma^2(\xi_t)\right) dt + \sigma(\xi_t) dW_t$$

and

$$\begin{aligned} E^s[S_t] &= E^x[e^{X_t}] \\ &= E\left[\exp\left(\int_0^t \mu(\xi_s) ds\right)\right] e^x. \end{aligned}$$

Proof. We will start by applying the infinitesimal generator from proposition 4.1 to the function $v(t,x,i) = E^{(x,i)}[e^{X_t}]$ like so

$$L_{X,\xi}[v](t,x,i) = \lim_{s \searrow 0} \frac{1}{s} E^{(x,i)}[v(t, X_s, \xi_s) - v(t,x,i)]$$

$$
\begin{aligned}
&= \lim_{s \searrow 0} \frac{1}{s} E^{(x,i)} \left[E^{(X_s, \xi_s)}[e^{X_t}] - E^{(x,i)}[e^{X_t}] \right] \\
&= \lim_{s \searrow 0} \frac{1}{s} E^{(x,i)} \left[E^{(x,i)}[e^{X_{t+s}} \mid \mathcal{F}_s] - E^{(x,i)}[e^{X_t} \mid \mathcal{F}_s] \right] \\
&= \lim_{s \searrow 0} \frac{1}{s} E^{(x,i)} \left[e^{X_{t+s}} - e^{X_t} \right] \\
&= \lim_{s \searrow 0} \frac{v(t+s,x,i) - v(t,x,i)}{s} \\
&= \frac{\partial v}{\partial t}(t,x,i).
\end{aligned}
$$

Now we make the observation that $\frac{\partial v}{\partial x} = v$ as shown below.

$$
\begin{aligned}
\frac{\partial v}{\partial x} &= \lim_{h \to 0} \frac{E^{(x+h,i)}[e^{X_t}] - E^{(x,i)}[e^{X_t}]}{h} \\
&= \lim_{h \to 0} \frac{E^{(x,i)}[e^{X_t+h} - e^{X_t}]}{h} \\
&= \lim_{h \to 0} \frac{e^h - 1}{h} E^{(x,i)}[e^{X_t}] \\
&= E^{(x,i)}[e^{X_t}] = v.
\end{aligned}
$$

Combining the two, we get

$$
\begin{aligned}
&\frac{\partial v}{\partial t}(t,x,i) = L_{(X,\xi)}[v](x,i) \\
&= \left(\mu(i) - \frac{1}{2}\sigma^2(i) \right) v_x(t,x,i) + \frac{1}{2}\sigma^2(i) v_{xx}(t,x,i) + \sum_{j=1}^{n} q_{ij} v(t,x,j) \\
&= \mu(i) v(t,x,i) + \sum_{j=1}^{n} q_{ij} v(t,x,j).
\end{aligned}
$$

Putting in matrix form, $\mathbf{V}' = (Q+U)\mathbf{V} \quad v(0,x,i) = e^x$ where $\mathbf{V}$ is the vector whose i-th element is $v(t,x,i)$ and U is the diagonal matrix whose i-th diagonal element is $\mu(i)$. Comparing this to proposition 2.1, it becomes clear that

$$
v(t,x,i) = E^{(x,i)} \left[\exp\left(\int_0^t \mu(\xi_s) ds \right) \right] e^x
$$

and the proof is complete. □

Now, from Proposition 5.1 we have that

$$E^{(s,i)}[e^{-rt}(S_t - K)] = e^{-rt}E^i\left[\exp\left(\int_0^t \mu(\xi_s)ds\right)\right]s - e^{-rt}K.$$

From Proposition 2.1 we have that

$$E^i\left[\exp\left(\int_0^t \mu(\xi_s)ds\right)\right] = M(t,i).$$

We define $M(t,i)$ as the i-th element of the vector

$$\mathbf{M}(t) = \sum_{i=1}^{n} C_i \mathbf{V_i} e^{\lambda_i t}$$

where $\mathbf{V}_i$ and λ_i are the i-th corresponding eigenvector and eigenvalue of $Q+U$ with U being the diagonal matrix whos i-th diagonal element is $\mu(i)$. The C_i's make up n arbitrary constants completely determined by the initial condition $\mathbf{M}(0) = 1$.

Combing the above, we get

$$E^{(s,i)}[e^{-rt}(S_t - K)] = \sum_{i=1}^{n}\left(sC_i\mathbf{V_i}e^{(\lambda_i - r)t}\right) - e^{-rt}K$$

and it becomes clear that if the real part of any of the eigenvalues of $Q + U$ are larger than r, then we expect the value of the contract to grow without bound and there is no optimal stopping strategy.

Conjecture 5.1. *If r is greater than the real part of all eigenvalues of $Q + U$ and $r \geq \max\{\mu(1), \ldots, \mu(n)\}$ then there exists an optimal stopping time for*

$$V(x,i) = \sup_{\tau} E^{(x,i)}[e^{-r\tau}(e^{X_\tau} - K)\mathbb{1}_{\{\tau<\infty\}}]$$

which is of the form

$$\tau = \inf\{t \geq 0 : V(X_t, \xi_t) = (e^{X_t} - K)\}$$

We will show that this stopping time is equivalent to having a series of thresholds, above which it is optimal to exercise the contract.

Lemma 5.1. *Assuming a stopping time of the form given in conjecture 5.1 there exists a series of thresholds,* $\{b(1), \ldots, b(n)\}$ *forming a region* $C = \{(x,i) : x < b(i)\}$ *referred to as the continuation region and a region* $D = \{(x,i) : x \geq b(i)\}$ *referred to as the stopping region such that the entry time* $\tau_D = \inf\{t \geq 0 : (X_t, \xi_t) \in D\} = \inf\{t \geq 0 : X_t \geq b(\xi_t)\}$ *is the optimal stopping time:*

$$V(x,i) = E^{(x,i)}\left[e^{-r\tau_D}(e^{X_{\tau_D}} - K)\mathbb{1}_{\{\tau_D<\infty\}}\right]$$

Modulo relabeling Markov states, we can and will assume $b(1) < \cdots < b(n)$.

Proof. We will show that $V(x,i) - \phi(x)$, where $\phi(x) = (e^x - K)$, is a decreasing function on $\mathbb{R}$. From this we conclude that so long as there is a value above which $V(x,i) = \phi(x)$ for all i, then there are unique thresholds depending on the Markov state above which $V(x,i) = \phi(x)$ and the lemma is proven. Let $x, x+\delta \in \mathbb{R}$ and let τ be the optimal stopping time when $X_0 = x$. We need the fact that $r \geq \max\{\mu(1), \ldots, \mu(n)\}$ implies that $e^{-rt}e^{X_t}$ is a supermartingale, which follows directly from proposition 5.1.

$$\begin{aligned}
V(j, x+\delta) &= E^{(x+\delta,j)}\left[e^{-r\tau}\phi(X_\tau)\right] \\
&= E^{(x,j)}\left[e^{-r\tau}\left(e^{X_\tau+\delta} - K\right)\right] \\
&= E^{(x,j)}\left[e^{-r\tau}\left(e^{X_\tau} - K + (e^\delta - 1)e^{X_\tau}\right)\right] \\
&= E^{(x,j)}\left[e^{-r\tau}(e^{X_\tau} - K)\right] + (e^\delta - 1)E^{(x,j)}\left[e^{-r\tau}e^{X_\tau}\right] \\
&= V(x,j) + (e^\delta - 1)E^{(x,j)}\left[e^{-r\tau}e^{X_\tau}\right] \\
&\leq V(x,j) + (e^\delta - 1)\liminf_{t\to\infty} E^{(x,j)}\left[e^{-r(t\wedge\tau)}e^{X_{t\wedge\tau}}\right] && (13) \\
&\leq V(x,j) + (e^\delta - 1)e^x && (14) \\
&= V(x,j) + \phi(x+\delta) - \phi(x).
\end{aligned}$$

In (13) Fatou's lemma is applied. Finally (14) is valid by Doobs optional sampling theorem since the discounted stock process $e^{-rt}e^{X_t}$ is a supermartingale and $t \wedge \tau$ is a bounded stopping time. Notice that no assumptions are made about the finiteness of τ.

The ordering of the n thresholds would involve a simple renaming of the Markov states to achieve descending order and the proof is complete. □

Next, the techniques from [5] will be used to show that the payoff function $V(x,i)$ solves the Dirichlet problem

$$L_{(X,\xi)}[V](x,i) = rV(x,i) \quad in \quad C \tag{15}$$

$$V(x,i) = e^x - K \quad in \quad D.$$

$C = \{(x,i) : V(x,i) > (e^x - K)^+\}$ and is referred to as the continuation region and D is the its complement in $\mathbb{R} \times \{1, 2, \cdots, n\}$, and is referred to as the stopping region. The entrance time into D is defined by $\tau_D = \inf\{t \geq 0 : (X_t, \xi_t) \in D\}$. Here, no assumptions are made as to the finiteness of τ_D.

To verify (15), we first define a "killed" process $\tilde{X}_t$. Let T be a killing time, with rate of killing defined as follows

$$\begin{aligned} r(\xi_t) &= \lim_{\delta \searrow 0} \frac{P(t < T \leq t + \delta \mid \{\xi_u : 0 \leq u < t\}, T > t)}{\delta} \\ &= \frac{f(t \mid \{\xi_u : 0 \leq u < t\})}{1 - F(t \mid \{\xi_u : 0 \leq u < t\})} \end{aligned}$$

with $F(t \mid \{\xi_u : 0 \leq u < t\}) = P(T \leq t \mid \{\xi_u : 0 \leq u < t\})$ and $f(t \mid \{\xi_u : 0 \leq u < t\})$ its derivative. Now if $r(\xi_t)$ is integrated from 0 to t, we get

$$\begin{aligned} \int_0^t r(\xi_s) ds &= \int_0^t \frac{f(s \mid \{\xi_u : 0 \leq u < s\})}{1 - F(s \mid \{\xi_u : 0 \leq u < s\})} ds \\ &= -\ln\Big(1 - F(s \mid \{\xi_u : 0 \leq u < s\})\Big) \end{aligned}$$

giving,

$$\begin{aligned} P(T \leq t \mid \{\xi_u : 0 \leq u < t\}) &= 1 - e^{-\int_0^t r(\xi_s) ds} \\ P(t \leq T \mid \{\xi_u : 0 \leq u < t\}) &= e^{-\int_0^t r(\xi_s) ds}. \end{aligned} \tag{16}$$

Define the killed process

$$\tilde{X}_t := \begin{cases} X_t & t < T \\ \Delta & t \geq T \end{cases},$$

where Δ is referred to as the cemetery (hence the terminology of killed process and killing time T). Define a function $\phi(\Delta) = 0$ and observe that from (16) we have

$$\begin{aligned} E[\phi(\tilde{X}_t)] &= E[\phi(X_t) \mathbb{1}_{t<T}] \\ &= E\Big[E[\phi(X_t) \mathbb{1}_{t<T} \mid \mathcal{F}_t]\Big] \\ &= E\Big[\phi(X_t) E[\mathbb{1}_{t<T} \mid \mathcal{F}_t]\Big] \\ &= E\Big[\phi(X_t) e^{-\int_0^t r(\xi_s) ds}\Big]. \end{aligned}$$

The above provides a convenient way to write the value function

$$V(x,i) = E\left[\phi(\tilde{X}_{\tau_D})\right]$$
$$\phi(x) = (e^x - K).$$

Next, the strong Markov property will be used to show that $L_{(X,\xi)}[F](x,i) = rF(x,i)$.

Choose $(x,i) \in C$ and a bounded open set $U \subset C$ and define $\sigma = \inf\{t : (X_t, \xi_t) \notin U\}$ when (X_t, ξ_t) starts at (x,i). Notice that $\sigma \le \tau_D$.

$$\begin{aligned} E^{(x,i)}[V(\tilde{X}_\sigma, \xi_\sigma)] &= E^{(x,i)}\left[E^{(\tilde{X}_\sigma,\xi_\sigma)}[\phi(\tilde{X}_{\tau_D})]\right] \\ &= E^{(x,i)}\left[E^{(x,i)}[\phi(\tilde{X}_{\tau_D}) \circ \theta_\sigma \mid \mathcal{F}_\sigma]\right] \\ &= E^{(x,i)}\left[E^{(x,i)}[\phi(\tilde{X}_{\sigma+\tau_D\circ\theta_\sigma}) \mid \mathcal{F}_\sigma]\right] \\ &= E^{(x,i)}\left[E^{(x,i)}[\phi(\tilde{X}_{\tau_D}) \mid \mathcal{F}_\sigma]\right] \\ &= E^{(x,i)}[\phi(\tilde{X}_{\tau_D})] \\ &= V(x,i). \end{aligned}$$

Thus the characteristic operator is identically zero:

$$\lim_{U \searrow (x,i)} \frac{E^{(x,i)}[V(\tilde{X}_\sigma, \xi_\sigma)] - V(x,i)}{E[\sigma]} = 0.$$

Since the characteristic and infinitesimal operator coincide on the domain of the infinitesimal operator, we have

$$L_{(\tilde{X},\xi)}[V](x,i) = 0 \quad for \quad (x,i) \in C.$$

The infinitesimal generator of the killed process is given by

$$L_{(\tilde{X},\xi)}[V](x,i) = L_{(X,\xi)}[V](x,i) - rV(x,i).$$

Applying Proposition 4.1 leads to the system of ODEs below

$$\left(\mu(i) - \frac{1}{2}\sigma^2(i)\right) f_x(x,i) + \frac{1}{2}\sigma^2(i) f_{xx}(x,i) + \sum_{j=1}^{n} q_{ij} f(x,j) = rf(x,i)$$
$$for \quad x < b_i \tag{17}$$
$$f(x,i) = e^x - K$$
$$for \quad x \ge b_i.$$

In solving the above system of ODEs we will start by looking at the system when $x < b_1$ to have a system with n unknown functions which can be written in matrix form

$$\mathcal{S}\boldsymbol{f}_x + \frac{1}{2}\Sigma \boldsymbol{f}_{xx} + (Q - R)\boldsymbol{f} = 0, \tag{18}$$

where Σ and R are the diagonal matrices whose i-th diagonal elements are $\sigma^2(i)$ and r respectively and Q is the infinitesimal generating matrix of the Markov chain. $\mathcal{S} = U - \frac{1}{2}\Sigma$ where U is the diagonal matrix whose i-th diagonal element is $\mu(i)$. $\boldsymbol{f}$ is the vector whose i-th element is $f(x,i)$. In the standard way, we seek a solution of the form $f(x,i) = g(i)e^{-\lambda x}$ leading to

$$(Q - R)\boldsymbol{g} - \lambda \mathcal{S}\boldsymbol{g} + \frac{1}{2}\lambda^2 \Sigma \boldsymbol{g} = 0$$

with $\boldsymbol{g}$ being the vector whose i-th element is $g(i)$. This method was presented in ([3], p. 2065) for solving the "quadratic eigenvalue" problem. Multiplying the above equation on the left by $2\Sigma^{-1}$ and reformulating as a system of equations yields,

$$\begin{cases} \lambda g &= h \\ \lambda h &= 2\Sigma^{-1}\mathcal{S}h - 2\Sigma^{-1}(Q - R)g, \end{cases}$$

which can be written as a the standard linear eigenvalue problem

$$\begin{pmatrix} 0 & I \\ -2\Sigma^{-1}(Q - R) & 2\Sigma^{-1}R - I \end{pmatrix} \begin{pmatrix} g \\ h \end{pmatrix} = \lambda \begin{pmatrix} g \\ h \end{pmatrix}.$$

As stated in [3], there are exactly n eigenvalues with a positive real part and n with a negative real part. For $x < b_1$ we will only consider the n values of λ with negative real part since we do not want the solution to (18), sought in the form $f(x,i) = g(i)e^{-\lambda x}$, to grow unbounded below b_1. If $\boldsymbol{g}_i$ and λ_i for $i = 1, \ldots, n$ all solve the above eigenvalue problem, we arrive at part of the solution to (17).

$$f(x, j) = \sum_{i=1}^{n} \omega_i g_i(j) e^{-\lambda_i x} \quad for \quad x < b_1 \quad j = 1, \ldots, n.$$

In particular, we have an entire solution for $f(x,1)$:

$$f(x,1) = \begin{cases} \sum_{i=1}^{n} \omega_i g_i(1) e^{-\lambda_i x} & ; x < b_1 \\ e^x - K & ; x \geq b_1. \end{cases}$$

Now the region $b_1 \leq x < b_2$ will be considered. In this region, $f(x,1) = e^x - K$, thus the size of the ODE system reduces by one. Let Q_1 denote the matrix Q with the first row and column removed and let $\tilde{Q}_1$ be matrix (vector in this case) composed of the first column of Q with the first row removed. Let $\mathcal{S}_1$, R_1, and Σ_1 be defined as the corresponding matrix with the first row and column removed and let $\boldsymbol{g}_1$ be the vector $\boldsymbol{g}$ with the first element removed. We arrive at the following matrix form of the $n-1$ dimensional system of ODEs

$$(Q_1 - R_1)\boldsymbol{g}_1 - \lambda \mathcal{S}_1 \boldsymbol{g}_1 + \frac{1}{2}\lambda^2 \Sigma \boldsymbol{g}_1 + \tilde{Q}_1(e^x - K) = 0 \tag{19}$$

for $b_1 \leq x < b_2$.

A particular solution to the above system of ODEs is of the form $B_1 + C_1 e^x$ where B_1 and C_1 are $n-1$ dimensional vectors. The values of B_1 and C_1 are completely determined by the ODEs. Let $B_1(j)$ and $C_1(j)$ represent the j-th component of the corresponding vector. The solution to the homogeneous equation,

$$(Q_1 - R_1)\boldsymbol{g}_1 - \lambda \mathcal{S}_1 \boldsymbol{g}_1 + \frac{1}{2}\lambda^2 \Sigma \boldsymbol{g}_1 = 0 \quad for \quad b_1 \leq x < b_2,$$

is solved precisely like (18). Since the region $b_1 \leq x < b_2$ is bounded, we will consider all $2(n-1)$ eigenvalues and eigenvectors of the homogeneous ODE. We arrive at another part of the solution to (17) and a complete solution for $f(x,2)$.

$$f(x,j) = \sum_{i=1}^{n} \omega_i^1 g_i^1(j) e^{-\lambda_i^1 x} + B_1(j-1) + C_1(j-1)e^x$$

for $b_1 \leq x < b_2$ and $j = 2, \ldots, n$, with

$$f(x,2) = \begin{cases} \sum_{i=1}^{n} \omega_i g_i(2) e^{-\lambda_i x} & ; b_1 > x \\ \sum_{i=1}^{2(n-1)} \omega_i^1 g_i^1(2) e^{-\lambda_i^1 x} + B_1(1) + C_1(1)e^x & ; b_1 \leq x < b_2 \\ e^x - K & ; x \geq b_2 \end{cases} .$$

Continuing on in the same manner we arrive at the region $b_k \leq x < b_{k+1}$. In this region, $f(x,j) = e^x - K$ for $j = 1, \ldots, k$, thus the size of the ODE system is reduced by k. Let Q_k denote the matrix Q with the first k rows and columns removed and let $\tilde{Q}_k$ be matrix composed of the first k columns of Q with the first k rows removed. Let $\boldsymbol{\Phi}_k$ be vector of size k where each component is $e^x - K$. Let $\mathcal{S}_k$, R_k, and Σ_k be defined as the corresponding matrix with the first k rows and columns removed and let $\boldsymbol{g}_k$ be the vector $\boldsymbol{g}$ with the first k elements removed. We arrive at the following matrix form of the $n-k$ dimensional system of ODEs,

$$(Q_k - R_k)\boldsymbol{g}_k - \lambda \mathcal{S}_k \boldsymbol{g}_k + \frac{1}{2}\lambda^2 \Sigma \boldsymbol{g}_k + \tilde{Q}_k \boldsymbol{\Phi}_k = 0 \quad for \quad b_k \leq x < b_{k+1}$$

solved the same way as (19). We arrive at another part of the solution to (17) and a complete solution for $f(x, k+1)$.

$$f(x, j) = \sum_{i=1}^{n} \omega_i^k g_i^k(j) e^{-\lambda_i^k x} + B_k(j-k) + C_k(j-k)e^x$$

for $b_k \leq x < b_{k+1}$ and $j = k+1, \ldots, n$ with

$$f(x, k+1) = \begin{cases} \sum_{i=1}^{n} \omega_i g_i(k+1) e^{-\lambda_i x} & ; x < b_1 \\ \sum_{i=1}^{2(n-1)} \omega_i^1 g_i^1(k+1) e^{-\lambda_i^1 x} + B_1(k) + C_1(k)e^x & ; b_1 \leq x < b_2 \\ \vdots & ; \vdots \\ \sum_{i=1}^{2(n-k)} \omega_i^k g_i^k(k+1) e^{-\lambda_i^k x} + B_k(1) + C_k(1)e^x & ; b_k \leq x < b_{k+1} \\ e^x - K & ; b_{k+1} \leq x. \end{cases}$$

This procedure is continued up to the point when $k = n-1$. Here, the system of ODEs becomes one ODE and the last function $f(x, n)$ is found in its entirety.

For the first function $f(x, 1)$ there are n unknown weights ω_i. The next function $f(x, 2)$ has $2(n-1)$ unknown weights, until the last function $f(x, n)$ contributes only 2 unknown weights. There are also n unknown boundary values b_i. Overall there are $2(1 + \cdots + n) = n(n+1)$ unknown parameters that need to be determined. Here, we assume that the function is C^1 everywhere. At this point, this assumption may seem restrictive, however, it will later be proven that the solution derived from this smoothness assumption is indeed the optimal solution. Assuming that $f(x, 1)$ is continuous and differentiable at b_1 will result in 2 conditions for $f(x, 1)$, 4 for $f(x, 2)$ with its two boundaries, and finally $2n$ for $f(x, n)$ and its n boundaries. In total, there are $n(n+1)$ conditions to be satisfied. We see that the $n(n+1)$ unknown parameters, including all weights and the n unknown boundaries, are completely determined by imposing the smoothness assumption.

Theorem 5.1 (Optimality)**.** *Suppose that the thresholds* $\ln(K) < b_1 < \cdots < b_n$ *have been found such that the unique solution to*

$$\left(\mu(i) - \frac{1}{2}\sigma^2(i)\right) f_x(x, i) + \frac{1}{2}\sigma^2(i) f_{xx}(x, i) + \sum_{j=1}^{n} q_{ij} f(x, j) = r f(x, i)$$

$$for \quad x < b_i$$

$$f(x, i) = \phi(x)$$

$$for \quad x \geq b_i.$$

is C^1 on its domain and bounded on C. Further, suppose the following assumptions hold

1. $f(x,i) \geq e^x - K$ *for all* (x,i)
2. r *is greater than the real part of all eigenvalues of* $Q + U$
3. $r > \max\{\mu(1), \ldots, \mu(n)\}$
4. $r > \max_{1\leq i\leq n}\left\{\frac{M(i)+\mu(i)e^{b(i)}}{e^{b(i)}-K}\right\}$ *where* $M(i) = \sum_{j=1}^{i-1} q_{ij}K$

Then the solution $f(x,i)$ and the stopping time $\tau_D = \{t : X_t \geq b(\xi_t)\}$ correspond to the value function

$$V(x,i) = \sup_{\tau} E^{(x,i)}[e^{-r\tau}(e^{X_\tau} - K)\mathbb{1}_{\{\tau<\infty\}}]$$

and its optimal stopping time, i.e.,

$$f(x,i) = V(x,i) = E^{(x,i)}[e^{-r\tau_D}(e^{X_{\tau_D}} - K)\mathbb{1}_{\{\tau_D<\infty\}}]$$

Proof. We will start by looking at the process $e^{-rt}f(X_t,\xi_t)$. Since by definition $f(x,i)$ is twice differentiable everywhere except when $x = b_i$ for $i = 1,2,\cdots,n$ where X_t spends zero time, we can apply the generalized Ito-formula (12) to it to get

$$e^{-rt}f(X_t,\xi_t) = f(X_0,\xi_0) + \int_0^t e^{-rs}\left(L_{(X,\xi)}[f](X_s,\xi_s) - rf(X_s,\xi_s)\right)ds + Martingale. \tag{20}$$

To show optimality we need the following

Proposition 5.2. *The inequality for the function defined below is true for all* (x,i)

$$\begin{aligned}\Phi(x,i) &:= L_{(X,\xi)}[f](x,i) - rf(x,i)\\ &= \left(\mu(i) - \frac{1}{2}\sigma^2(i)\right)f_x(x,i) + \frac{1}{2}\sigma^2(i)f_{xx}(x,i) - rf(x,i) + \sum_{j=1}^{n} q_{ij}f(x,j)\\ &\leq 0.\end{aligned}$$

Proof. Using the hypothesis $f(x,i) \geq \phi(x)$ we first show that $f(x,i) - \phi(x)$ is decreasing in x:

$$
\begin{aligned}
V(j, x-\delta) &= E^{(x-\delta,j)}\left[e^{-r(t\wedge\tau_D)} f(X_{t\wedge\tau_D}, \xi_{t\wedge\tau_D})\right] && (21)\\
&= E^{(x-\delta,j)}\left[e^{-r\tau_D} f(X_{\tau_D}, \xi_{\tau_D})\right] && (22)\\
&\geq E^{(x-\delta,j)}\left[e^{-r\tau_D}\phi(X_{\tau_D})\right]\\
&= E^{(x,j)}\left[e^{-r\tau_D}\left(e^{X_{\tau_D}-\delta} - K\right)\right]\\
&= E^{(x,j)}\left[e^{-r\tau_D}\left(e^{X_{\tau_D}} - K + (e^{-\delta}-1)e^{X_{\tau_D}}\right)\right]\\
&= E^{(x,j)}\left[e^{-r\tau_D}(e^{X_{\tau_D}} - K)\right] + (e^{-\delta}-1)E^{(x,j)}\left[e^{-r\tau_D}e^{X_{\tau_D}}\right]\\
&= V(x,j) + (e^{-\delta}-1)E^{(x,j)}\left[e^{-r\tau_D}e^{X_{\tau_D}}\right]\\
&\geq V(x,j) + (e^{-\delta}-1)\liminf_{t\to\infty} E^{(x,j)}\left[e^{-r(t\wedge\tau_D)}e^{X_{t\wedge\tau_D}}\right] && (23)\\
&\geq V(x,j) + (e^{-\delta}-1)e^{x} && (24)\\
&= V(x,j) + \phi(x-\delta) - \phi(x).
\end{aligned}
$$

(21) is valid since $e^{-r(t\wedge\tau_D)} f(X_{t\wedge\tau_D}, \xi_{t\wedge\tau_D})$ is a martingale due to $\Phi(x,i) = 0$ for $(x,i) \in C$. We note that when $X_0 = x - \delta$ at any time t less than τ_D, X_t is still in region C despite the fact that τ_D is found as if $X_0 = x$. In (23) and (22) Fatou's lemma is applied. Finally (24) is valid by Doobs optional sampling theorem since the discounted stock process $e^{-rt}e^{X_t}$ is a supermartingale and $t \wedge \tau_D$ is a bounded stopping time.

The above yields

$$
\begin{aligned}
f(x,i) - \phi(x) &\leq \lim_{x\to-\infty} f(x,i) - \phi(x)\\
&= \lim_{x\to-\infty} \sum_{j=1}^{n} \omega_j g_j(i) e^{-\lambda_j x} - e^x - K\\
&= K
\end{aligned}
$$

due to the fact that λ_i are chosen to be negative.

For $(x,i) \in C$, we have $\Phi(x,i) = 0$ by construction. For $(x,i) \in D$ we have $f(x,i) = \phi(x)$ and

$$
\begin{aligned}
\Phi(x,i) &:= L_{(X,\xi)}[f](x,i) - rf(x,i)\\
&= \left(\mu(i) - \frac{1}{2}\sigma^2(i)\right) f_x(x,i) + \frac{1}{2}\sigma^2(i) f_{xx}(x,i) - rf(x,i) + \sum_{j=1}^{n} q_{ij} f(x,j)
\end{aligned}
$$

$$
\begin{aligned}
&= e^x(\mu(i) - r) + rK + \sum_{\substack{j=1 \\ j \neq i}}^{n} q_{ij}(f(x,j) - \phi(x)) \\
&= e^x(\mu(i) - r) + rK + \sum_{j=1}^{i-1} q_{ij}(f(x,j) - \phi(x)) \\
&\leq e^x(\mu(i) - r) + rK + \sum_{j=1}^{i-1} q_{ij}K \\
&= -r(e^x - K) + \mu(i)e^x + M(i) \\
&\leq -\frac{M(i) + \mu(i)e^x}{e^x - K}(e^x - K) + \mu(i)e^x + M(i) \\
&= 0.
\end{aligned}
$$

The last line follows from the fact that $\frac{M+\mu(i)e^x}{e^x-K}$ is decreasing function and thus

$$r > \frac{M + \mu(n)e^{b(i)}}{e^{b(i)} - K} \geq \frac{M + \mu(n)e^x}{e^x - K}$$

Now we have that $\Phi(x,i) \leq 0$ as desired. □

By proposition 5.2, we have

$$e^{-rt}\phi(X_t, \xi_t) \leq e^{-rt} f(X_t, \xi_t) \leq f(x,i) + Martingale.$$

Therefore for any stopping time τ

$$E^{(x,i)}[e^{-r(t\wedge\tau)}\phi(X_{t\wedge\tau}, \xi_{t\wedge\tau})] \leq f(x,i)$$

and by Fatou's lemma

$$E^{(x,i)}[e^{-r\tau}\phi(X_\tau, \xi_\tau)] \leq f(x,i).$$

Finally we observe that $f(X_{\tau_D}, \xi_{\tau_D}) = \phi(X_{\tau_D}, \xi_{\tau_D})$ and that $e^{-r(t\wedge\tau)} f(X_{t\wedge\tau_D}, \xi_{t\wedge\tau_D})$ is a martingale due to $\Phi(x,i) = 0$ for $(x,i) \in C$. Furthermore, it is a uniformly integrable martingale since $x \leq b(n)$ when the process is stopped at τ_D and $f(x,i)$ is bounded below $b(n)$. Using Doobs optional sampling theorem we get

$$f(x,i) = E^{(x,i)}[e^{-r\tau_D} f(X_{\tau_D}, \xi_{\tau_D})] = E^{(x,i)}[e^{-r\tau_D}\phi(X_{\tau_D}, \xi_{\tau_D})]$$

and optimality is proven. □

References

[1] S. N. Ethier and T. G. Kurtz, *Markov Processes Characterization and Convergence*, University Press, USA (1996).

[2] X. Guo, Q. Zhang, Optimal Selling Rules in a Regime Switching Model, *Transactions on Automatic Control*, **50**, No. 9 (2005), 1450-1455, **doi:** 10.1109/TAC.2005.854657.

[3] A. Jobert, L.C.G. Rogers, Option Pricing with Markov-Modulated Dynamics, *SIAM J. Control and Optimization*, **44**, No. 6 (2006), 2063-2078, **doi:** 10.1137/050623279.

[4] F.C. Klebaner, *Introduction to Stochastic Calculus with Applications*, Imperial College Press, UK (2012).

[5] G. Peskir, A. Shiryaev, *Optimal Stopping and Free-Boundary Problems*, Birkhauser Verlag, USA (2006).

[6] P. Protter, *Stochastic Integration and Differential Equations*, Springer Verlag, USA (2010).