THE UNIVERSITY OF
WARWICK

Original citation:
Deshpande, Amogh. (2014) Sufficient stochastic maximum principle for the optimal control of semi-Markov modulated jump-diffusion with application to financial optimization. Stochastic Analysis and Applications, Volume 32 (Number 6). pp. 911-933. ISSN 1532-9356

Permanent WRAP url:
http://wrap.warwick.ac.uk/62025

Publisher's statement:
The Version of Record of this manuscript has been published and is available in Stochastic Analysis and Applications, 2014
http://www.tandfonline.com/doi/abs/10.1080/07362994.2014.945038#.VGszevmsV8F

warwick**publications**wrap
highlight your research

http://wrap.warwick.ac.uk

Sufficient stochastic maximum principle for the optimal control of semi-Markov modulated jump-diffusion with application to Financial optimization.

Amogh Deshpande [*]

Abstract

Paper forthcoming in Stochastic Analysis and Applications

The finite state semi-Markov process is a generalization over the Markov chain in which the sojourn time distribution is any general distribution. In this article we provide a sufficient stochastic maximum principle for the optimal control of a semi-Markov modulated jump-diffusion process in which the drift, diffusion and the jump kernel of the jump-diffusion process is modulated by a semi-Markov process. We also connect the sufficient stochastic maximum principle with the dynamic programming equation. We apply our results to finite horizon risk-sensitive control portfolio optimization problem and to a quadratic loss minimization problem.

Keywords: semi-Markov modulated jump diffusions, sufficient stochastic maximum principle, dynamic programming, risk-sensitive control, quadratic loss-minimization.

AMS subject classification 93E20; 60H30;46N10.

---

[*]Department of Statistics, University of Warwick, Coventry, CV47AL,UK. Email: addeshpa@gmail.com

1 Introduction

The stochastic maximum principle is a stochastic version of the Pontryagin maximum principle which states that the any optimal control must satisfy a system of forward-backward stochastic differential equations, called the optimality system, and should maximize a functional, called the Hamiltonian. The converse indeed is true and gives the sufficient stochastic maximum principle. In this article we will derive sufficient stochastic maximum principle for a class of process called as the semi-Markov modulated jump-diffusion process. In this process the drift, the diffusion and the jump kernel term is modulated by an semi-Markov process.

An early investigation of stochastic maximum principle and its application to finance has been credited to Cadenillas and Karatzas [1]. Framstadt et al. [5] formulated the stochastic maximum principle for jump-diffusion process and applied it to a quadratic portfolio optimization problem. Their work has been partly generalized by Donnelly [4] who considered a Markov chain modulated diffusion process in which the drift and the diffusion term is modulated by a Markov chain. Zhang et al. [11] studied sufficient maximum principle of a process similar to that studied by Donnelly additionally with a jump term whose kernel is also modulated by a Markov chain. It can be noted that the Markov modulated process has been quite popular with its recent applications to finance for example Options pricing (Deshpande and Ghosh [3]) and references therein and to portfolio optimization refer Xhou and Yin [13]. However application of semi-Markov modulated process to portfolio optimization in which the portfolio wealth process is a semi-Markov modulated diffusion are not many, see for example Ghosh and Goswami [6]. Even so it appears that the sufficient maximum principle has not been formulated for the case of a semi-Markov modulated diffusion process with jumps and studied further in the context of quadratic portfolio optimization. Moreover, application of the sufficient stochastic maximum principle in the context of risk-sensitive control portfolio optimization with the portfolio wealth process following a semi-Markov modulated diffusion process has not been studied. This article aims to provide answers to these missing dots and connect them together. For the same reasons, alongwith providing a popular application of the sufficient stochastic maximum principle to a quadratic loss minimization problem when the portfolio wealth process follows a semi-Markov modulated jump-diffusion, we also provide an example of risk-sensitive portfolio optimization for the diffusion part of the said dynamics.

The article is organized as follows. In the next section we formally describe basic terminologies used in the article. In section 3 we detail the control problem that we are going to study. The

sufficient maximum principle is proven in Section 4. This is followed by establishing its connection with the dynamic programming. We conclude the article by illustrating its applications to risk-sensitive control optimization and to a quadratic loss minimization problem.

2 Mathematical Preliminaries

We adopt the following notations that are valid for the whole paper:

$\mathbb{R}$: the set of real numbers

r, M: any positive integer greater than 1.

$\mathcal{X} = \{1, ..., M\}$.

$\mathcal{C}^{1,2,1}([0, T] \times \mathbb{R}^r \times \mathcal{X} \times \mathbb{R}_+)$: denote the family of all functions on $[0, T] \times \mathbb{R}^r \times \mathcal{X} \times \mathbb{R}_+$ which are twice continuously differentiable in x and continuously differentiable in t and y.

v', A': the transpose of the vector (say)v and matrix say A respectively.

$||v||$: Euclidean norm of a vector v.

$|A|$: norm of a matrix A.

$tr(A)$: trace of a square matrix A.

$C_b^m(\mathbb{R}^r)$: Set of real m-times continuously differentiable functions which are bounded together with their derivatives upto the mth order.

We assume that the probability space $(\Omega, \mathcal{F}, \{\mathcal{F}(t)\}, \mathbb{P})$ is complete with filtration $\{\mathcal{F}(t)\}_{t \geq 0}$ and is right-continuous and $\mathcal{F}(0)$ contains all $\mathbb{P}$ null sets. Let $\{\theta(t)\}_{t \geq 0}$ be a semi-Markov process taking values in $\mathcal{X}$ with transition probability p_{ij} and conditional holding time distribution $F^h(t|i)$. Thus if $0 \leq t_0 \leq t_1 \leq ...$ are times when jumps occur, then

$$P(\theta(t_{n+1}) = j, t_{n+1} - t_n \leq t | \theta(t_n) = i) = p_{ij} F^h(t|i). \tag{2.1}$$

Matrix $[p_{ij}]_{\{i,j=1,...,M\}}$ is irreducible and for each i, $F^h(\cdot|i)$ has continuously differentiable and bounded density $f^h(\cdot|i)$. For a fixed t, let $n(t) \triangleq \max\{n : t_n \leq t\}$ and $Y(t) \triangleq t - t_{n(t)}$. Thus $Y(t)$ represents the amount of time the proess $\theta(t)$ is at the current state after the last jump. The process $(\theta(t), Y(t))$defined on $(\Omega, \mathcal{F}, \mathbb{P})$ is jointly Markov and the differential generator $\mathcal{L}$ given as follows (Chap.2, [7])

$$\mathcal{L}\phi(i, y) = \frac{d}{dy}\phi(i, y) + \frac{f^h(y|i)}{1 - F^h(y|i)} \sum_{j \neq i, j \in \mathcal{X}} p_{ij}[\phi(j, 0) - \phi(i, y)]. \tag{2.2}$$

for $\phi : \mathcal{X} \times \mathbb{R}_+ \to \mathbb{R}$ is C^1 function.

We first represent semi-Markov process $\theta(t)$ as a stochastic integral with respect to a Poisson

random measure. With that perspective in mind, embed $\mathcal{X}$ in $\mathbb{R}^M$ by identifying i with $e_i \in \mathbb{R}^M$. For $y \in [0, \infty)$ $i, j \in \mathcal{X}$, define

$$\lambda_{ij}(y) \;=\; p_{ij} \frac{f^h(y/i)}{1 - F^h(y/i)} \geq 0 \;\; \text{and} \;\; \forall \; i \neq j,$$

$$\lambda_{ii}(y) \;=\; - \sum_{j \in \mathcal{X}, j \neq i}^{M} \lambda_{ij}(y) \;\; \forall \; i \;\; \in \mathcal{X}.$$

For $i \neq j \in \mathcal{X}$, $y \in \mathbb{R}_+$ let $\Lambda_{ij}(y)$ be consecutive (with respect to lexicographic ordering on $\mathcal{X} \times \mathcal{X}$) left-closed, right-open intervals of the real line, each having length $\lambda_{ij}(y)$. Define the functions $\bar{h} : \mathcal{X} \times \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}^r$ and $\bar{g} : \mathcal{X} \times \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}_+$ by

$$\bar{h}(i, y, z) = \begin{cases} j - i & \text{if } z \in \Lambda_{ij}(y) \\ 0 & \text{otherwise} \end{cases}$$

$$\bar{g}(i, y, z) = \begin{cases} y & \text{if } z \in \Lambda_{ij}(y), j \neq i \\ 0 & \text{otherwise} \end{cases}$$

Let $\mathcal{M}(\mathbb{R}_+ \times \mathbb{R})$ be the set of all nonnegative integer-valued σ-finite measures on Borel σ-field of $(\mathbb{R}_+ \times \mathbb{R})$. The process $\{\tilde{\theta}(t), Y(t)\}$ is defined by the following stochastic integral equations:

$$\tilde{\theta}(t) = \tilde{\theta}(0) + \int_0^t \int_{\mathbb{R}} \bar{h}(\tilde{\theta}(u-), Y(u-), z) N_1(du, dz),$$

$$Y(t) = t - \int_0^t \int_{\mathbb{R}} \bar{g}(\tilde{\theta}(u-), Y(u-), z) N_1(du, dz),$$

(2.3)

where $N_1(dt, dz)$ is an $\mathcal{M}(\mathbb{R}_+ \times \mathbb{R})$-valued Poisson random measure with intensity $dt m(dz)$ independent of the $\mathcal{X}$-valued random variable $\tilde{\theta}(0)$, where $m(\cdot)$ is a Lebesgue measure on $\mathbb{R}$. As usual by definition $Y(t)$ represents the amount of time, process $\tilde{\theta}(t)$ is at the current state after the last jump. We define the corresponding compensated or centered one dimensional Poisson measure as $\tilde{N}_1(ds, dz) = N_1(ds, dz) - ds m(dz)$. It was shown in Theorem 2.1 of Ghosh and Goswami [6] that $\tilde{\theta}(t)$ is a semi-Markov process with transition probability matrix $[p_{ij}]_{\{i, j = 1, \ldots, M\}}$ with conditional holding time distributions $F^h(y|i)$. Since by definition $\theta(t)$ is also a semi-Markov process with transition probability matrix $[p_{ij}]_{\{i, j = 1, \ldots, M\}}$ with conditional holding time distributions $F^h(y|i)$ defined on the same underlying probability space, by equivalence, $\tilde{\theta}(t) = \theta(t)$ for $t \geq 0$.

Remark 2.1 The semi-Markov process with conditional density $f^h(y|i) = \tilde{\lambda}_i e^{-\tilde{\lambda}_i y}$ for some $\tilde{\lambda}_i > 0$, $i = 1, 2 \ldots, M$, is infact a Markov chain.

3 The control problem

Let $\mathcal{U} \subset \mathbb{R}^r$ be a closed subset. Let $\mathbb{B}_0$ be the family of Borel sets $\Gamma \subset \mathbb{R}^r$ whose closure $\bar{\Gamma}$ does not contain 0. For and Borel set $B \subset \Gamma$, one dimensional poisson random measure $N(t, B)$ counts the number of jumps on $[0, t]$ with values in B. For a predictable process $u : [0, T] \times \Omega \to \mathcal{U}$ with right continuous left limit paths, consider the controlled process X with given initial condition $X(0) = x \in \mathbb{R}^r$ given by

$$dX(t) = b(t, X(t), u(t), \theta(t))dt + \sigma(t, X(t), u(t), \theta(t))dW(t) + \int_\Gamma g(t, X(t), u(t), \theta(t)), \gamma)N(dt, d\gamma),$$
$$(3.1)$$

where $X(t) \in \mathbb{R}^r$ and $W(t) = (W_1(t), ..., W_r(t))$ is r-dimensional standard Brownian motion. The coefficients $b(\cdot, \cdot, \cdot, \cdot) : [0, T] \times \mathbb{R}^r \times \mathcal{U} \times \mathcal{X} \to \mathbb{R}^r, \sigma(\cdot, \cdot, \cdot, \cdot) : [0, T] \times \mathbb{R}^r \times \mathcal{U} \times \mathcal{X} \to \mathbb{R}^r \times \mathbb{R}^r$ and $g(\cdot, \cdot, \cdot, \cdot, \cdot) : [0, T] \times \mathbb{R}^r \times \mathcal{U} \times \mathcal{X} \times \Gamma \to \mathbb{R}^r$ and satisfy the following conditions,

Assumption (A1)

(At most linear growth) There exists a constant $C_1 < \infty$ for any $i \in \mathcal{X}$ such that
$|\sigma(t, x, u, i)|^2 + ||b(t, x, u, i)||^2 + \int_\mathbb{R} ||g(t, x, u, i, \gamma)||^2 \lambda(d\gamma) \leq C_1(1 + ||x||^2)$

(Lipschitz continuity) There exists a constant $C_2 < \infty$ for any $i \in \mathcal{X}$ such that
$|\sigma(t, x, u, i) - \sigma(t, y, u, i)|^2 + ||b(t, x, u, i) - b(t, y, u, i)||^2 + \int_\Gamma ||g(t, x, u, i, \gamma) - g(t, y, u, i, \gamma)||^2 \lambda(d\gamma) \leq C_2 ||x - y||^2 \ \forall x, y \in \mathbb{R}^r$.

Then $X(t)$ is a unique cadlag adapted solution given by (3.1) refer Theorem 1.19 of [10].

Define $a(t, x, u, i) = \sigma(t, x, u, i)\sigma'(t, x, u, i)$ is a $\mathbb{R}^{r \times r}$ matrix and $a_{kl}(t, x, u, i)$ is the $(k, l)^{th}$ element of the matrix a while $b_k(t, x, u, i)$ is the kth element of the vector $b(t, x, u, i)$. We assume that $N(\cdot, \cdot), N_1(\cdot, \cdot)$ and θ_0, W_t, X_0 defined on $(\Omega, \mathcal{F}, \mathbb{P})$ are independent. For future use we define the compensated Poisson measure $\tilde{N}(dt, d\gamma) = N(dt, d\gamma) - \lambda\pi(d\gamma)dt$, where $\pi(\cdot)$ is the jump distribution (is a probability measure) and $0 < \lambda < \infty$ is the jump rate such that $\int_\Gamma \min(||\gamma||^2, 1)\lambda(d\gamma) < \infty$.

Consider the performance criterion

$$J^u(x, i, y) = E^{x,i,y}[\int_0^T f_1(t, X(t), u(t), \theta(t), Y(t))dt + f_2(X(T), \theta(T), Y(T))], \quad (3.2)$$

where $f_1 : [0, T] \times \mathbb{R}^r \times \mathcal{U} \times \mathcal{X} \times \mathbb{R}_+ \to \mathbb{R}$ is continuous and $f_2 : \mathbb{R}^r \times \mathcal{X} \times \mathbb{R}_+ \to \mathbb{R}$ is concave. We say that the admissible class of controls $u \in \mathcal{A}(T)$ if

$$E^{x,i,y}\left[\int_0^T |f_1(t, X(t), u(t), \theta(t), Y(t))|dt + f_2(X(T), \theta(T), Y(T))]\right] < \infty.$$

The problem is to maximize J^u over all $u \in \mathcal{A}(T)$ i.e. we seek $\hat{u} \in \mathcal{A}(T)$ such that

$$J^{\hat{u}}(x, i, y) = \sup_{u \in \mathcal{A}(T)} J^u(x, i, y), \tag{3.3}$$

where $\hat{u}$ is an optimal control.

Define a Hamiltonian $\mathcal{H} : [0, T] \times \mathbb{R}^r \times \mathcal{U} \times \mathcal{X} \times \mathbb{R}_+ \times \mathbb{R}^r \times \mathbb{R}^{r \times r} \times \mathbb{R}^r \to \mathbb{R}$ by,

$$
\begin{aligned}
\mathcal{H}(t, x, u, i, y, p, q, \eta) \quad := \quad & f_1(t, x, u, i, y) + \left(b^{'}(t, x, u, i) - \int_\Gamma g^{'}(t, x, u, i, \gamma) \pi(d\gamma) \right) p + tr(\sigma^{'}(t, x, u, i) q) \\
& + \quad \left(\int_\Gamma g^{'}(t, x, u, i, \gamma) \pi(d\gamma) \right) \eta. \tag{3.4}
\end{aligned}
$$

We assume that the Hamiltonian $\mathcal{H}$ is differentiable with respect to x. The adjoint equation corresponding to u and X^u in the unknown adapted processes $p(t) \in \mathbb{R}^r$, $q(t) \in \mathbb{R}^{r \times r}$, $\eta : \mathbb{R}_+ \times \mathbb{R}^r - \{0\} \to \mathbb{R}^r$ and $\tilde{\eta}(t, z) = (\eta^{(1)}(t, z), ..., \eta^{(r)}(t, z))^{'}$, where $\tilde{\eta}^{(n)}(t, z) \in \mathbb{R}^{r \times r}$ for each $n = 1, 2, ..., r$, is the backward stochastic differential equation (BSDE),

$$
\begin{aligned}
dp(t) \quad = \quad & -\nabla_x \mathcal{H}(t, X(t), u(t), \theta(t), p(t), q(t), \eta(t, \gamma)) dt + q^{'}(t) dW(t) + \int_\Gamma \eta(t, \gamma) \tilde{N}(dt, d\gamma) \\
& + \quad \int_\mathbb{R} \tilde{\eta}(t, z) \tilde{N}_1(dt, dz), \\
p(T) \quad = \quad & \nabla_x f_2(X(T), \theta(T), Y(T)). \quad a.s. \tag{3.5}
\end{aligned}
$$

We have assumed that $\mathcal{H}$ is differentiable with respect to $x = X(t)$ and is denoted as $\nabla_x \mathcal{H}(t, X(t), u(t), \theta(t), p(t), q(t), \eta(t, \gamma))$. As per Remark 2.1, for the special case where the semi-Markov process has exponential holding time distribution, we would have (3.5) to be a BSDE with Markov chain switching. For this special case, Cohen and Elliott [2] have provided conditions for uniqueness of the solution. However, corresponding uniqueness result for the semi-Markov modulated BSDE as in (3.5) seems not available in the literature. Since this paper concerns sufficient conditions, we will assume ad hoc that a solution to this BSDE exists and is unique.

Remark 3.1 Notice that there are jumps in the adjoint equation (3.5) attributed to jumps in the semi-Markov process $\theta(t)$. This is because the drift, the diffusion and the jump kernel of the process $X(t)$ is modulated by a semi-Markov process. Also note that the unknown process $\tilde{\eta}(t, z)$ in the adjoint equations (3.5) does not appear in the Hamiltonian (3.4).

4 Sufficient Stochastic Maximum principle

In this section we state and prove the sufficient stochastic maximum principle.

Theorem 4.1(Sufficient Maximum principle) Let $\hat{u} \in \mathcal{A}(T)$ with corresponding solution $\hat{X} \triangleq X^{\hat{u}}$.

Suppose there exists a solution $(\hat{p}(t), \hat{q}(t), \hat{\eta}(t, \gamma), \hat{\tilde{\eta}}(t, z))$ of the adjoint equation (3.5) satisfying

$$E \int_0^T ||\left(\sigma(t, \hat{X}(t), \theta(t)) - \sigma(t, X^u(t), \theta(t))\right)' \hat{p}(t)||^2 dt < \infty \qquad (4.1)$$

$$E \int_0^T ||\hat{q}'(t)\left(\hat{X}(t) - X^u(t)\right)||^2 dt < \infty \qquad (4.2)$$

$$E \int_0^T ||(\hat{X}(t) - X^u(t))' \hat{\eta}(t, \gamma)||^2 \pi(d\gamma) dt < \infty \qquad (4.3)$$

$$E \int_0^T |\left(\hat{X}(t) - X^u(t)\right)' \hat{\tilde{\eta}}(t, z)|^2 m(dz) dt < \infty. \qquad (4.4)$$

for all admissible controls $u \in \mathcal{A}(T)$. If we further suppose that

1.

$$\mathcal{H}(t, \hat{X}(t), \hat{u}(t), \theta(t), Y(t), \hat{p}(t), \hat{q}(t), \hat{\eta}(t, \cdot)) = \sup_{u \in \mathcal{A}(T)} \mathcal{H}(t, \hat{X}(t), u(t), \theta(t), Y(t), \hat{p}(t), \hat{q}(t), \hat{\eta}(t, \cdot)). (4.5)$$

2. for each fixed pair $(t, i, y) \in ([0, T] \times \mathcal{X} \times \mathbb{R}_+)$, $\hat{\mathcal{H}}(x) := \sup_{u \in \mathcal{A}(T)} \mathcal{H}(t, x, u, i, y, \hat{p}(t), \hat{q}(t), \hat{\eta}(t, \cdot))$

exists and is a concave function of x. Then $\hat{u}$ is an optimal control.

Proof   Fix $u \in \mathcal{A}(T)$ with corresponding solution $X = X^u$. For sake of brevity we would hence-forth represent $(t, \hat{X}(t-), \hat{u}(t-), \theta(t-), Y(t-))$ by $(t, \hat{X}(t-))$ and $(t, X(t-), u(t-), \theta(t-), Y(t-))$ by $(t, X(t-))$. Then,

$$J(\hat{u}) - J(u) = E\left(\int_0^T \left(f_1(t, \hat{X}(t)) - f_1(t, X(t))\right) dt + f_2(\hat{X}(T), \theta(T), Y(T)) - f_2(X(T), \theta(T), Y(T)) \right).$$

By use of concavity of $f_2(\cdot, i, y)$ we have for each $i \in \mathcal{X}$, $y \in \mathbb{R}_+$ and (3.5) to obtain the inequalities,

$$E\left(f_2(\hat{X}(T), \theta(T), Y(T)) - f_2(X(T), \theta(T), Y(T)) \right) \geq E\left((\hat{X}(T) - X(T))' \nabla_x f_2(\hat{X}(T), \theta(T), Y(T)) \right)$$

$$\geq E\left((\hat{X}(T) - X(T))' \hat{p}(T) \right).$$

which gives

$$J(\hat{u}) - J(u) \geq E \int_0^T \left(f_1(t, \hat{X}(t)) - f_1(t, X(t))\right) dt + E\left((\hat{X}(T) - X(T))' \hat{p}(T) \right). \qquad (4.6)$$

7

We now expand the above equation (4.6) term by term. For the first term in this equation we use the definition of $\mathcal{H}$ as in (3.4) to obtain

$$
E \int_0^T \Big(f_1(t, \hat{X}(t)) - f_1(t, X(t)) \Big) dt
$$

$$
= \quad E \int_0^T \Big(\mathcal{H}(t, \hat{X}(t), \hat{u}(t), \theta(t), \hat{p}(t), \hat{q}(t), \hat{\eta}(t, \gamma))
$$

$$
- \quad \mathcal{H}(t, X(t), u(t), \theta(t), p(t), q(t), \eta(t, \gamma)) \Big) dt
$$

$$
- \quad E \int_0^T \Big[\Big(b(t, \hat{X}(t)) - b(t, X(t))
$$

$$
- \quad \int_\Gamma \Big(g(t, \hat{X}(t-), \hat{u}(t-), \theta(t-), \gamma) - g(t, X(t-), u(t-), \theta(t-), \gamma) \Big) \pi(d\gamma) \Big) \hat{p}(t)
$$

$$
+ \quad tr\Big((\sigma(t, \hat{X}(t)) - \sigma(t, X(t)))' \hat{q}(t) \Big)
$$

$$
+ \quad \int_\Gamma (g(t, \hat{X}(t-), \hat{u}(t-), \theta(t-), \gamma) - g(t, X(t-), u(t-), \theta(t-), \gamma))' \eta(t, \gamma) \pi(d\gamma) \Big] dt.
$$

$$
(4.7)
$$

To expand the second term on the right hand side of (4.6) we begin by applying the integration by parts formula to get,

$$
(\hat{X}(T) - X(T))' \hat{p}(T) \quad = \quad \int_0^T (\hat{X}(t) - X(t))' d\hat{p}(t)
$$

$$
+ \quad \int_0^T \hat{p}'(t) d(\hat{X}(t) - X(t)) + [\hat{X} - X, \hat{p}](T).
$$

Substitute for X, $\hat{X}$ and $\hat{p}$ from (3.1) and (3.5) to obtain,

$$
(\hat{X}(T) - X(T))' \hat{p}(T)
$$
$$
= \int_0^T (\hat{X}(t) - X(t))' \Big(-\nabla_x \mathcal{H}(t, \hat{X}(t), \hat{u}(t), \hat{p}(t), \hat{q}(t), \hat{\eta}(t, \gamma)) dt + \hat{q}'(t) dW(t)
$$
$$
+ \int_\Gamma \hat{\eta}(t, \gamma) \tilde{N}(dt, d\gamma) + \int_\mathbb{R} \hat{\tilde{\eta}}(t, z) \tilde{N}_1(dt, dz) \Big)
$$
$$
+ \int_0^T \hat{p}'(t) \Big\{ \Big(\big(b(t, \hat{X}(t)) - b(t, X(t)) \big) - \int_\Gamma \big(g(t, \hat{X}(t), \hat{u}(t-), \theta(t-), \gamma)
$$
$$
- g(t, X(t-), u(t-), \theta(t-), \gamma) \big) \pi(d\gamma) \Big) dt
$$
$$
+ \Big(\sigma(t, \hat{X}(t)) - \sigma(t, X(t)) \Big)' dW(t)
$$
$$
+ \int_\Gamma \Big(g(t, \hat{X}(t-), \hat{u}(t-), \theta(t-), \gamma) - g(t, X(t-), u(t-), \theta(t-), \gamma) \Big) \tilde{N}(dt, d\gamma) \Big\}
$$
$$
+ \int_0^T \Big[tr \Big(\hat{q}'(t) \Big(\sigma(t, \hat{X}(t)) - \sigma(t, X(t)) \Big) \Big)
$$
$$
+ \int_\Gamma \Big(g(t, \hat{X}(t), \hat{u}(t-), \theta(t-), \gamma) - g(t, X(t), u(t-), \theta(t-), \gamma) \Big)' \eta(t, \gamma) \pi(d\gamma) \Big] dt.
$$

Due to integrability conditions (4.1)-(4.4), the integral with respect to the Brownian motion and the Poisson random measure are square integrable martingales which are null at the origin. Thus taking expectations we obtain

$$
E \Big((\hat{X}(T) - X(T))' \hat{p}(T) \Big)
$$
$$
= \int_0^T (\hat{X}(t) - X(t))' \Big(-\nabla_x \mathcal{H}(t, \hat{X}(t), \hat{u}(t), \hat{p}(t), \hat{q}(t), \hat{\eta}(t, \gamma)) \Big) dt
$$
$$
+ \int_0^T \Big[\hat{p}'(t) \Big(b(t, \hat{X}(t)) - b(t, X(t)) - \int_\Gamma \big(g(t, \hat{X}(t-), \hat{u}(t-), \theta(t-), \gamma)
$$
$$
- g(t, X(t), u(t-), \theta(t-), \gamma) \big) \pi(d\gamma) \Big)
$$
$$
+ \int_0^T tr \Big(\hat{q}'(t) (\sigma(t, \hat{X}(t)) - \sigma(t, X(t))) \Big)
$$
$$
+ \int_\Gamma \Big(\big(g(t, \hat{X}(t-), \theta(t-), u(t-), \gamma) - g(t, X(t-), \theta(t-), u(t-), \gamma) \big)' \eta(t, \gamma) \Big) \pi(d\gamma) \Big] dt.
$$

Substitute the last equation and (4.7) into the inequality (4.6) to find after cancellation that

$$
J(\hat{u}) - J(u) \geq E \int_0^T \Big(\mathcal{H}(t, \hat{X}(t), \hat{u}(t), \theta(t), \hat{p}(t), \hat{q}(t), \hat{\eta}(t, \gamma)) - \mathcal{H}(t, X(t), u(t), \theta(t), p(t), q(t), \eta(t, \gamma))
$$
$$
- (\hat{X}(t) - X(t))' \nabla_x \mathcal{H}(t, \hat{X}(t), \hat{u}(t), \theta(t), \hat{p}(t), \hat{q}(t), \hat{\eta}(t, \gamma)) \Big) dt. \tag{4.8}
$$

We can show that the integrand on the RHS of (4.8) is non-negative a.s. for each $t \in [0, T]$ by fixing the state of the semi-Markov process and then using the assumed concavity of $\hat{\mathcal{H}}(x)$, we apply the argument in Framstad et al. [5] . This gives $J(\hat{u}) \geq J(u)$ and $\hat{u}$ is an optimal control. ∎

5 Connection to the Dynamic programming

We show the connection between the stochastic maximum principle and dynamic programming principle for the semi-Markov modulated regime switching jump diffusion. This tantamounts to explicitly showing connection between the value function $V(t, x, i, y)$ of the control problem and the adjoint processes $p(t), q(t), \eta(t, \gamma)$ and $\tilde{\eta}(t, z)$. In order to apply the dynamic programming principle we put the problem into a Markovian framework by defining

$$J^u(t, x, i, y) \triangleq E^{X(t)=x, \theta(t)=i, Y(t)=y}\Big[\int_t^T f_1(t, X(t), u(t), \theta(t), Y(t))dt + f_2(X(T), \theta(T), Y(T))\Big]. \tag{5.1}$$

and put

$$V(t, x, i, y) = \sup_{u \in \mathcal{A}(T)} J^u(t, x, i, y) \quad \forall \ (t, x, i, y) \in [0, T] \times \mathbb{R}^r \times \mathcal{X} \times \mathbb{R}_+. \tag{5.2}$$

Theorem 5.1 *Assume that $V(\cdot, \cdot, i, \cdot) \in \mathcal{C}^{1,3,1}([0, T] \times \mathbb{R}^r \times \mathcal{X} \times \mathbb{R}_+)$ for each $i, j \in \mathcal{X}$ and that there exists an optimal Markov control $\hat{u}(t, x, i, y)$ for (5.2), with the corresponding solution $\hat{X} = X^{(\hat{u})}$. Define*

$$p_k(t) \quad \triangleq \quad \frac{\partial V}{\partial x_k}(t, \hat{X}(t), \theta(t), Y(t)). \tag{5.3}$$

$$q_{kl}(t) \quad \triangleq \quad \sum_{i=1}^r \sigma_{il}(t, \hat{X}(t), \hat{u}(t), \theta(t)) \frac{\partial^2 V}{\partial x_i \partial x_k}(t, \hat{X}(t), \theta(t), Y(t)). \tag{5.4}$$

$$\eta^{(k)}(t, \gamma) \quad \triangleq \quad \frac{\partial V}{\partial x_k}(t, \hat{X}(t), j, Y(t)) - \frac{\partial V}{\partial x_k}(t, \hat{X}(t), i, Y(t)). \tag{5.5}$$

$$\tilde{\eta}^{(k)}(t, z) \quad \triangleq \quad \frac{\partial V}{\partial x_k}(t, \hat{X}(t-), \theta(t-) + \bar{h}(\theta(t-), Y(t-), z), Y(t-) - \bar{g}(\theta(t-), Y(t-), z))$$
$$- \quad \frac{\partial V}{\partial x_k}(t, \hat{X}(t-), \theta(t-), Y(t-)). \tag{5.6}$$

for each $(k, l = 1, ..., r)$. Also we assume that the coefficients $b(t, x, u, i)$, $\sigma(t, x, u, i)$ and $g(t, x, u, i, \gamma)$ belong to $C_b^1(\mathbb{R}^r)$. Then $p(t), q(t), \eta(t, \gamma)$ and $\tilde{\eta}(t, z)$ solves the adjoint equation (3.5).

10

We prove this theorem by using the following Ito's formula.

Theorem 5.2 Suppose r dimensional process $X(t) = (X_1(t), ..., X_r(t))$ or $\{X_g(t)\}$ indexed by $(g = 1, 2, ..., r)$ satisfies the following equation,

$$dX_g(t) = b_g(t, X(t), u(t), \theta(t))dt + \sum_{m=1}^{r} \sigma_{gm}(t, X(t), u(t), \theta(t))dW_m(t) + \int_{\Gamma} g_g(t, X(t-), u(t), \theta(t-), \gamma)N(dt, d\gamma).$$

for some $X(0) = x_0 \in \mathbb{R}^r$ $a.s.$. Further let us assume that the coefficients b, σ, g satisfies the conditions of Assumption (A1).

Let $V(\cdot, \cdot, i, \cdot) \in C^{1,3,1}([0, T] \times \mathbb{R}^r \times \mathcal{X} \times \mathbb{R}_+)$. Then the generalized Ito's formula is given by

$$V(t, X(t), \theta(t), Y(t)) - V(t, x, \theta, y) = \int_0^t GV(s, X(s), \theta(s), Y(s))ds$$

$$+ \int_0^t (\nabla_x V(s, X(s), \theta(s), Y(s)))' \sigma(s, X(s), \theta(s))dW(s)$$

$$+ \int_0^t \int_{\Gamma} [V(s, X(s-) + g(s, X(s-), u(s), \theta(s-), \gamma), \theta(s-), Y(s-))$$

$$- V(s, X(s-), \theta(s-), Y(s-))]\tilde{N}(ds, d\gamma)$$

$$+ \int_0^t \int_{\mathbb{R}} [V(s, X(s-), \theta(s-) + \bar{h}(\theta(s-), Y(s-), z), Y(s-) - \bar{g}(\theta(s-), Y(s-), z))$$

$$- V(s, X(s-), \theta(s-), Y(s-))]\tilde{N}_1(ds, dz),$$

where the local martingale terms are explicitly defined as

$$dM_1(t) \triangleq (\nabla_x V(t, X(t), \theta(t), Y(t)))' \sigma(t, X(t), u(t), \theta(t))dW_t,$$

$$dM_2(t) \triangleq \int_{\Gamma} [V(t, X(t-) + g(t, X(t-), u(t), \theta(t-), \gamma), \theta(t-), Y(t-)) - V(t, X(t-), \theta(t-), Y(t-))]\tilde{N}(dt, d\gamma),$$

$$dM_3(t) \triangleq \int_{\mathbb{R}} [V\Big(t, X(t-), \theta(t-) + \bar{h}(\theta(t-), Y(t-), z), Y(t-) - \bar{g}(\theta(t-), Y(t-), z)\Big)$$

$$- V(t, X(t-), \theta(t-), Y(t-))]\tilde{N}_1(dt, dz),$$

for

$$
\begin{aligned}
GV(t,x,i,y) \;=\;\; & \frac{\partial V(t,x,i,y)}{\partial t} \\
+ \;\; & \frac{1}{2}\sum_{g,l=1}^{r} a_{gl}(t,x,i)\frac{\partial V(t,x,i,y)}{\partial x_g \partial x_l} \\
+ \;\; & \sum_{g=1}^{r} b_g(t,x,i)\frac{\partial V(t,x,i,y)}{\partial x_g} \\
+ \;\; & \frac{\partial V(t,x,i,y)}{\partial y} \\
+ \;\; & \frac{f^h(y|i)}{1-F^h(y|i)}\sum_{j\neq i,j\in\mathcal{X},i=1}^{M} p_{ij}[V(t,x,j,0)-V(t,x,i,y)] \\
+ \;\; & \lambda\int_{\Gamma}(V(t,x+g(t,x,i,\gamma),i,y)-V(t,x,i,y))\pi(d\gamma),
\end{aligned}
$$

$\forall\, t\,\in\,[0,T]\,, x\in\mathbb{R}^r, (i=1,....,M),\; y\in\mathbb{R}_+.$

Proof For details refer to Theorem 5.1 in Ikeda and Watanabe [8]. ∎

Proof of Theorem 5.1 From the standard theory of the Dynamic programming the following HJB equation holds:

$$
\frac{\partial V}{\partial t}(t,x,i,y)+\sup_{u\in\mathcal{U}}\{f_1(t,x,u,i,y)+\mathcal{A}^u V(t,x,i,y)\}=0,
$$
$$
V(T,x,i,y)=f_2(x,i,y).
$$

where $\mathcal{A}^u$ is the infinitesimal generator and the supremum is attained by $\hat{u}(t,x,i,y)$. Define

$$
F(t,x,u,i,y)=f_1(t,x,u,i,y)+\frac{\partial V}{\partial t}(t,x,i,y)+\mathcal{A}^u V(t,x,i,y).
$$

We assume that f_1 is differentiable w.r.t to x. We use the Ito's formula as described in Theorem 5.2 to get,

$$
\begin{aligned}
F(t,x,u,i,y) \;=\;\; & f_1(t,x,u,i,y)+\frac{\partial V}{\partial t}(t,x,i,y) \\
+ \;\; & \sum_{k=1}^{r}\frac{\partial V}{\partial x_k}(t,x,i,y)b_k(t,x,u,i)+\frac{1}{2}\sum_{k=1}^{r}\sum_{l=1}^{r}\frac{\partial^2 V}{\partial x_k \partial x_l}(t,x,i,y)\sum_{i=1}^{r}\sigma_{ki}(t,x,u,i)\sigma_{li}(t,x,u,i) \\
+ \;\; & \sum_{j\neq i,i=1}^{M}\frac{p_{ij}f^h(y|i)}{1-F^h(y|i)}(V(t,x,j,0)-V(t,x,i,y))+\frac{\partial V}{\partial y}(t,x,i,y) \\
+ \;\; & \lambda\int_{\Gamma}(V(t,x+g(t,x,u,i,\gamma),i,y)-V(t,x,i,y))\pi(d\gamma). \qquad (5.7)
\end{aligned}
$$

Differentiate $F(t, x, \hat{u}(t, x, i, y), i, y)$ with respect to x_g and evaluate at $x = \hat{X}(t)$, $i = \theta(t)$ and $y = Y(t)$, we get,

$$
\begin{aligned}
0 \;=\;& \frac{\partial f_1}{\partial x_g}(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t), Y(t)) \\
&+ \frac{\partial^2 V}{\partial x_g \partial t}(t, \hat{X}(t), \theta(t), Y(t)) + \sum_{k=1}^{r} \frac{\partial^2 V}{\partial x_g \partial x_k}(t, \hat{X}(t), \theta(t), Y(t)) b_k(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t)) \\
&+ \sum_{k=1}^{r} \frac{\partial V}{\partial x_k}(t, \hat{X}(t), \theta(t), Y(t)) \frac{\partial b_k}{\partial x_g}(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t)) \\
&+ \frac{1}{2} \sum_{k=1}^{r} \sum_{l=1}^{r} \frac{\partial^3 V}{\partial x_g \partial x_k \partial x_l}(t, \hat{X}(t), \theta(t), Y(t)) \\
&\times \sum_{i=1}^{r} \sigma_{k,i}(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t)) \sigma_{l,i}(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t)) \\
&+ \frac{1}{2} \sum_{k=1}^{r} \sum_{l=1}^{r} \frac{\partial^2 V}{\partial x_k \partial x_l}(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t), Y(t)) \\
&\times \frac{\partial}{\partial x_g} \sum_{i=1}^{r} \sigma_{k,i}(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t)) \sigma_{l,i}(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t)) \\
&+ \sum_{j \neq i, j \in \mathcal{X}}^{M} \frac{p_{ij} f^h(y|i)}{1 - F^h(y|i)} \left(\frac{\partial V}{\partial x_g}(t, \hat{X}(t), j, 0) - \frac{\partial V}{\partial x_g}(t, \hat{X}(t), i, y) \right) \\
&+ \lambda \int_{\Gamma} \left(\frac{\partial V}{\partial x_g}(t, \hat{X}(t) + g(t, \hat{X}(t), \theta(t), \gamma), \theta(t), Y(t)) - \frac{\partial V}{\partial x_g}(t, \hat{X}(t), \theta(t), Y(t)) \right) \pi(d\gamma). \qquad (5.8)
\end{aligned}
$$

Next define, $Y_g = \frac{\partial V}{\partial x_g}(t, \hat{X}(t), \theta(t), Y(t))$ for $(g = 1, ..., r)$. By Ito's formula (Theorem 5.2) we obtain the dynamics of $Y_g(t)$ as follows,

$$
\begin{aligned}
dY_g(t) &= \Bigg\{ \frac{\partial^2 V}{\partial x_g \partial t}(t, \hat{X}(t), \theta(t), Y(t)) + \sum_{k=1}^{r} \frac{\partial^2 V}{\partial x_g \partial x_k}(t, \hat{X}(t), \theta(t), Y(t)) b_k(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t)) \\
&+ \frac{1}{2} \sum_{k=1}^{r} \sum_{l=1}^{r} \frac{\partial^3 V}{\partial x_g \partial x_k \partial x_l}(t, \hat{X}(t), \theta(t), Y(t)) \\
&\times \sum_{i=1}^{r} \sigma_{ki}(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t)) \times \sigma_{li}(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t)) \\
&+ \sum_{j \neq i, j=1}^{M} \frac{p_{ij} f^h(y|i)}{1 - F^h(y|i)} \Big(\frac{\partial V}{\partial x_g}(t, \hat{X}(t), j, 0) - \frac{\partial V}{\partial x_g}(t, \hat{X}(t), i, y) \Big) \\
&+ \lambda \int_{\Gamma} \Big(\frac{\partial V}{\partial x_g}(t, \hat{X}(t) + g(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t), \gamma), \theta(t), Y(t)) \\
&- \frac{\partial V}{\partial x_g}(t, \hat{X}(t), \theta(t), Y(t)) \Big) \pi(d\gamma) \Bigg\} dt \\
&+ \sum_{k=1}^{r} \frac{\partial^2 V}{\partial x_g \partial x_k}(t, \hat{X}(t), \theta(t), Y(t)) \sum_{j=1}^{r} \sigma_{kj}(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t)) dW_j(t) \\
&+ \int_{\Gamma} \Bigg\{ \frac{\partial V}{\partial x_g}(t, \hat{X}(t-) + g(t, \hat{X}(t-), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t-), \gamma), \theta(t-), Y(t-)) \\
&- \frac{\partial V}{\partial x_g}(t, \hat{X}(t-), \theta(t-), Y(t-)) \Bigg\} \tilde{N}(dt, d\gamma) \\
&+ \int_{\mathbb{R}} \Bigg\{ \frac{\partial V}{\partial x_g}((t, X(t-), \theta(t-) + \bar{h}(\theta(t-), Y(t-), z), Y(t-) - \bar{g}(\theta(t-), Y(t-), z))) \\
&- \frac{\partial V}{\partial x_g}(t, \hat{X}(t-), \theta(t-), Y(t-)) \Bigg\} \tilde{N}_1(dt, dz).
\end{aligned}
$$

We substitute $\frac{\partial^2 V}{\partial x_g \partial t}$ from (5.8) to get,

$$
\begin{aligned}
dY_g(t) \;=\; & -\frac{\partial f_1}{\partial x_g}(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t), Y(t))) \\
& - \sum_{k=1}^{r} \frac{\partial V}{\partial x_k}(t, \hat{X}(t), \theta(t), Y(t)) \frac{\partial b_k}{\partial x_g}(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t)) \\
& - \frac{1}{2} \sum_{k=1}^{r} \sum_{l=1}^{r} \frac{\partial^2 V}{\partial x_k \partial x_l}(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t), Y(t)) \\
& \times \frac{\partial}{\partial x_g} \Big(\sum_{k=1}^{r} \sigma_{ki}(t, \hat{X}(t), \theta(t)) \sigma_{li}(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t)) \Big) \\
& + \sum_{k=1}^{r} \frac{\partial^2 V}{\partial x_g \partial x_k}(t, \hat{X}(t), \theta(t), Y(t)) \sum_{j=1}^{r} \sigma_{kj}(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t)) dW_j(t) \\
& + \int_{\Gamma} \Big\{ \big(\frac{\partial V}{\partial x_g}(t, \hat{X}(t-) + g(t, X(t-), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t-), \gamma), \theta(t-), Y(t-)) \\
& - \frac{\partial V}{\partial x_g}(t, \hat{X}(t-), \theta(t-), Y(t-))) \Big\} \tilde{N}(dt, d\gamma) \\
& + \int_{\mathbb{R}} \Big\{ \frac{\partial V}{\partial x_g}((t, X(t-), \theta(t-) + \bar{h}(\theta(t-), Y(t-), z), Y(t-) - \bar{g}(\theta(t-), Y(t-), z))) \\
& - \frac{\partial V}{\partial x_g}(t, \hat{X}(t-), \theta(t-), Y(t-)) \Big\} \tilde{N}_1(dt, dz). \quad (5.9)
\end{aligned}
$$

We have the following identity,

$$
\begin{aligned}
& \frac{1}{2} \sum_{k=1}^{r} \sum_{l=1}^{r} \frac{\partial^2 V}{\partial x_k \partial x_l}(t, \hat{X}(t), \theta(t), Y(t)) \\
& \times \frac{\partial}{\partial x_g} \Big(\sum_{i=1}^{r} \sigma_{ki}(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t)) \sigma_{li}(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t)) \Big) \\
& = \sum_{k=1}^{r} \sum_{l=1}^{r} \sum_{i=1}^{r} \sigma_{il}(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t)) \frac{\partial^2 V}{\partial x_i \partial x_k}(t, \hat{X}(t), \theta(t), Y(t)) \\
& \times \frac{\partial \sigma_{kl}}{\partial x_g}(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t)). \quad (5.10)
\end{aligned}
$$

Next, from (3.4) we obtain,

$$
\frac{\partial \mathcal{H}}{\partial x_g}(t, X(t), u(t), \theta(t), Y(t), p(t), q(t), \eta(t, \gamma))
$$

$$
= \frac{\partial f_1}{\partial x_g}(t, \hat{X}(t), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t), Y(t))
$$

$$
+ \sum_{i=1}^{r} \left(\frac{\partial b_i}{\partial x_g}(t, \hat{X}(t-), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t-)) \right.
$$

$$
\left. - \int_{\Gamma} \frac{\partial g_i}{\partial x_g}(t, X(t-), \hat{u}(t, \hat{X}(t), \theta(t), Y(t)), \theta(t-), \gamma) \pi(d\gamma) \right) p_i(t) + tr(\frac{\partial \sigma'(t, x, \hat{u}, \theta(t))}{\partial x_g} q)
$$

$$
+ \sum_{i=1}^{r} \int_{\Gamma} \frac{\partial g_i}{\partial x_g}(t, X(t-), \theta(t-), \gamma) \pi(d\gamma)(\eta_i^{(g)}(t, \gamma)). \tag{5.11}
$$

We also note that

$$
tr(\frac{\partial \sigma'(t, x, u, i)}{\partial x_g} q) = \sum_{l=1}^{r} [\frac{\partial \sigma'(t, x, u, i)}{\partial x_g} q]_{ll}
$$

$$
= \sum_{l=1}^{r} \sum_{k=1}^{r} q_{k,l} \frac{\partial \sigma_{kl}}{\partial x_g}(t, x, u, i).
$$

Substitute (5.3)-(5.6) and (5.11) gives,

$$
dY_g(t) = -\frac{\partial \mathcal{H}}{\partial x_g}(t, X(t), u(t), \theta(t), Y(t), p(t), q(t), \eta(t, \gamma))dt + \sum_{j=1}^{r} q_{gj}(t)dW_j(t)
$$

$$
+ \int_{\Gamma} \eta(t, \gamma)\tilde{N}(dt, d\gamma) + \int_{\mathbb{R}} \tilde{\eta}(t, z)\tilde{N}_1(dt, dz). \tag{5.12}
$$

Since $Y_g(t) = p_g(t)$ for each $g = 1, ..., r$, we have shown that $p(t), q(t), \eta(t, \gamma)$ and $\tilde{\eta}(t, z)$ solve the adjoint equation (3.5). ∎

6 Applications

We illustrate the theory developed towards applying it to some key financial wealth optimization problems. For an early motivation on applying sufficient maximum principle, we first consider wealth dynamics to follow semi-Markov modulated diffusion (no jumps case) and apply it towards the risk-sensitive control portfolio optimization problem. We follow it up by illustrating an application of semi-Markov modulated jump-diffusion wealth dynamics to a quadratic loss minimization problem. Unless otherwise stated, all the processes defined in this section are one dimensional.

Risk-sensitive control portfolio optimization Let us consider a financial market consisting

of two continuously traded securities namely the risk less bond and a stock. The dynamics of the riskless bond is known to follow

$$dS_0(t) = r(t, \theta(t-))S_0(t)dt \quad S_0(0) = 1.$$

where $r(t, \theta(t))$ is the risk-free interest rate at time t and is modulated by an underlying semi-Markov process as described earlier. The dynamics of the stock price is given as

$$dS_1(t) = S_1(t)[(\mu(t, \theta(t-)))dt + \sigma(t, \theta(t-))dW(t)],$$

where $(\mu(t, \theta(t-)))$ is the instantaneous expected rate of return and as usual $\sigma(t, \theta(t-))$ is the instantaneous volatility rate. The stock price process is thus driven by a 1-d Brownian motion. We denote the wealth of the investor to be $X(t) \in \mathbb{R}$ at time t. He holds $\theta_1(t)$ units of stock and $\theta_0(t) = 1 - \theta_1(t)$ units is held in the riskless bond market. From the self-financing principle (refer Karatzas and Shreve [9]), the wealth process follows the dynamics given as,

$$dX(t) = (r(t, \theta(t-))X(t) + h(t)\sigma(t, \theta(t-))\bar{m}(t, \theta(t-)))dt + h(t)\sigma(t, \theta(t-))dW(t) \quad X(0) = x,$$

where $h(t) = \theta_1(t)S_1(t)$, $\bar{m}(t, i) = \frac{\mu(t,i) - r(t,i)}{\sigma(t,i)} \geq 0$ and the variables $r(t, i), b(t, i)$ and $\sigma(t, i)$, and $\sigma^{-1}(t, i)$ for each $i \in \mathcal{X}$ are measurable and uniformly bounded in $t \in [0, T]$. Also $h(\cdot)$ occuring in the drift and diffusion term in above dynamics of $X(t)$ satisfies the following conditions

1. $E[\int_0^T h^2(t)dt] < \infty$
2. $E[\int_0^T |r(t, \theta(t-))X(t) + h(t)\sigma(t, \theta(t-))\bar{m}(t, \theta(t-))|dt + \int_0^T h^2(t)\sigma^2(t, \theta(t-))dt] < \infty$
3. The SDE for X has a unique strong solution.

These conditions on $h(\cdot)$ are needed in order to prevent doubling strategies which otherwise would yield arbitrary profit at time T for an investor.

In a classical risk-sensitive control optimization problem, the investor aims to maximize over some admissible class of portfolio $\mathcal{A}(T)$ the following risk-sensitive criterion given by

$$
\begin{aligned}
J(\hat{h}(\cdot), x) &= \max_{h \in \mathcal{A}(T)} \frac{1}{\gamma}\mathbb{E}[X(T)^\gamma | X(0) = x, \theta(0) = i, Y(0) = y], \quad \gamma \in (1, \infty) \\
&= -\min_{h \in \mathcal{A}(T)} \frac{1}{\gamma}\mathbb{E}[X(T)^\gamma | X(0) = x, \theta(0) = i, Y(0) = y],
\end{aligned}
$$

where the exogenous parameter γ is the usual risk-sensitive criterion that describes the risk attitude of an investor. Thus the optimal expected utility function depends on γ and is a generalization of the traditional stochastic control approach to utility optimization in the sense that now the degree of risk aversion of the investor is explicitly parameterized through γ rather than importing it in the

problem via an exogenous utility function. See Whittle [12] for general overview on risk-sensitive control optimization. We now use the sufficient maximum principle (Theorem 4.1). Set the control problem $u(t) \triangleq h(t)$.

The corresponding Hamiltonian (for the non-jump case)(3.4) becomes,

$$\mathcal{H}(t, x, u, i, p, q) = (r(t,i)x + u\sigma(t,i)\bar{m}(t,i))p + u\sigma(t,i)q.$$

The adjoint process (3.5) is given by

$$
\begin{aligned}
dp(t) &= -r(t, \theta(t-))p(t)dt + q(t)dW(t) + \int_{\mathbb{R}} \tilde{\eta}(t,z)\tilde{N}_1(dt, dz), \\
p(T) &= X(T)^{\gamma-1} \quad a.s..
\end{aligned}
$$
(6.1)

We need to determine $p(t), q(t)$ and $\eta(t, z)$ in (6.1). Going by the terminal condition $p(T)$ we observe that the adjoint process p is the first derivative of (x^γ). Hence we assume that $p(t)$ defined as,

$$p(t) = (X(t))^{\gamma-1} e^{\phi(t, \theta(t), Y(t))}.$$

where $\phi(T, \theta(T) = i, Y(T)) = 0$ $a.s.$ for each $i \in \{1, ..., M\}$. Using the Ito's formula we get,

$$
\begin{aligned}
\frac{dp(t)}{p(t)} &= \sum_{i=1}^M 1_{\theta(t-)=i} \Bigg((\gamma-1)\Bigg\{ (r(t, \theta(t-))) + \frac{u(t)\sigma(t, \theta(t-))\bar{m}(t, \theta(t-))}{X(t)} \Bigg) \\
&+ \frac{1}{2}(\gamma-1)(\gamma-2)\sigma^2(t, \theta(t-))\frac{u^2(t)}{X^2(t)} \\
&+ \phi_t(t, \theta(t-), y) + \phi_y(t, \theta(t-), y) + \frac{f^h(y|\theta(t-)=i)}{1 - F^h(y|\theta(t-)=i)} \sum_{j \neq i} p_{ij}(\phi(t, j, 0) - \phi(t, \theta(t-), y)) \Bigg\} dt \\
&+ (\gamma-1)\frac{u(t)}{X(t)}\sigma(t, \theta(t-))dW(t) \\
&+ \int_{\mathbb{R}} \Bigg(\phi(t, X(t-), \theta(t-) + \bar{h}(\theta(t-), Y(t-), z), Y(t-) - \bar{g}(\theta(t-), Y(t-), z)) \\
&- \phi(t, \theta(t-), Y(t-)) \Bigg) \tilde{N}_1(dt, dz).
\end{aligned}
$$
(6.2)

Comparing the coefficient of (6.2) with that in (6.1) we get

$$
\begin{aligned}
-r(t, \theta(t-)) &= \sum_{i=1}^M 1_{\theta(t-)=i} \Bigg((\gamma-1)\Bigg(r(t, \theta(t-)) + \frac{u(t)\sigma(t, \theta(t-))\bar{m}(t, i)}{X(t)} \Bigg) + \frac{1}{2}(\gamma-1)(\gamma-2)\frac{u^2(t)}{X^2(t)} \\
&+ \phi_t(t, \theta(t-), y) + \phi_y(t, \theta(t-), y) + \frac{f^h(y|i)}{1 - F^h(y|\theta(t-)=i)} \sum_{j \neq i} p_{ij}(\phi(t, j, 0) - \phi(t, \theta(t-), y)) \Bigg).
\end{aligned}
$$
(6.3)

$$q(t) = (\gamma - 1)\frac{u(t)}{X(t)}\sigma(t, \theta(t-))p(t). \tag{6.4}$$

$$\tilde{\eta}(t, z) = \bigg(\phi(t, \theta(t-) + \bar{h}(\theta(t-), Y(t-), z), Y(t-) - \bar{g}(\theta(t-), Y(t-), z))$$

$$- \phi(t, \theta(t-), Y(t-))\bigg)p(t). \tag{6.5}$$

Let $\hat{u} \in \mathcal{A}(T)$ be a candidate optimal control corresponding to the wealth process $\hat{X}$ and the adjoint triplet $(\hat{p}, \hat{q}, \hat{\eta})$, then from the Hamiltonian (3.4) for all $u \in \mathbb{R}$ we have

$$\mathcal{H}(t, \hat{X}(t), u, \theta(t), \hat{p}(t), \hat{q}(t)) = \bigg(r(t, \theta(t))\hat{X}(t) + u\sigma(t, \theta(t))\bar{m}(t, \theta(t))\bigg)\hat{p}(t) + u\sigma(t, \theta(t))\hat{q}(t). \tag{6.6}$$

As this is a linear function of u, we guess that the coefficient of u vanishes at optimality, which results in the equality

$$\bar{m}(t, \theta(t-))\hat{p}(t) + \hat{q}(t) = 0. \tag{6.7}$$

Substitute equation (6.7) in (6.4) to obtain the expression for the control as

$$\hat{u}(t) = \frac{\bar{m}(t, \theta(t-))}{(1 - \gamma)\sigma(t, \theta(t-))}\hat{X}(t). \tag{6.8}$$

We now aim to determine the explicit expression for $p(t)$ which is only possible if we can determine what $\phi(t, \theta(t), Y(t))$ is. We substitute $\hat{u}$ from above and input it in equation (6.3) to get

$$\begin{aligned}
0 = & \gamma r(t, \theta(t-)) - \bar{m}^2(t, \theta(t-)) + \frac{(2 - \gamma)}{(1 - \gamma)}\frac{\bar{m}^2(t, \theta(t-))}{2\sigma^2(t, \theta(t-))} \\
& + \phi_t(t, \theta(t-), y) + \phi_y(t, \theta(t-), y) + \frac{f^h(y|\theta(t-) = i)}{1 - F^h(y|\theta(t-) = i)}\sum_{i=1, j\neq i}^{M} p_{ij}(\phi(t, j, 0) - \phi(t, \theta(t-), y)).
\end{aligned} \tag{6.9}$$

with terminal boundary condition given as $\phi(T, \theta(T), Y(T)) = 0$ a.s. Consider the process

$$\tilde{\phi}(t, \theta(t), Y(t)) \triangleq E\bigg[\exp\bigg(\int_t^T \bigg\{\gamma r(s, \theta(s)) - \bar{m}^2(s, \theta(s)) + \frac{(2 - \gamma)}{(1 - \gamma)}\frac{\bar{m}^2(s, \theta(s))}{2\sigma^2(s, \theta(s))}\bigg\}ds\bigg)|\theta(t-) = i, Y(t-) = y\bigg]. \tag{6.10}$$

We aim to show that $\phi = \tilde{\phi}$. For the same we define the following martingale,

$$R(t) \triangleq E\bigg[\exp\bigg(\int_0^T \bigg\{\gamma r(s, \theta(s)) - \bar{m}^2(s, \theta(s)) + \frac{(2 - \gamma)}{(1 - \gamma)}\frac{\bar{m}^2(s, \theta(s))}{2\sigma^2(s, \theta(s))}\bigg\}ds\bigg)|\mathcal{F}_t^{\theta, y}\bigg], \tag{6.11}$$

19

where $\mathcal{F}_\tau^{\theta,y} \triangleq \sigma\{\theta(\tau), Y(\tau), \tau \in [0,t]\}$ augmented with $\mathbb{P}$ null sets is the filtration generated by the processes $\theta(t)$ and $Y(t)$. From the $\{\mathcal{F}_t^{\theta,y}\}$-martingale representation theorem, there exist $\{\mathcal{F}_t^{\theta,y}\}$-previsible, square integrable process $\nu(t,i,y)$ such that

$$R(t) = R(0) + \int_0^t \int_\mathbb{R} \nu(\tau, \theta(\tau-), Y(\tau-)) \tilde{N}_1(d\tau, dz). \tag{6.12}$$

By positivity of $R(t)$ we can define $\hat{\nu}(\tau, \theta(\tau-), Y(\tau-)) \triangleq (\nu(\tau, \theta(\tau-), Y(\tau-))) R^{-1}(\tau-)$ so that

$$R(t) = R(0) + \int_0^t \int_\mathbb{R} R(\tau-) \hat{\nu}(\tau, \theta(\tau-), Y(\tau-)) \tilde{N}_1(d\tau, dz). \tag{6.13}$$

From the definition of $\tilde{\phi}$ in (6.10) and the definition of R in (6.11) it is easy to see that we have the following relationship

$$R(t) = \tilde{\phi}(t, \theta(t), Y(t)) \exp\left\{ \int_0^t (\gamma r(s, \theta(s)) - \bar{m}^2(s, \theta(s)) + \frac{(2-\gamma)}{(1-\gamma)} \frac{\bar{m}^2(s, \theta(s))}{2\sigma^2(s, \theta(s))}) ds \right\},$$
$$\forall \ t \ \in \ [0,T]. \tag{6.14}$$

Using the Ito's expansion of $\tilde{\phi}(t, \theta(t), Y(t))$ to the RHS of (6.14) followed up by comparing it with martingale representation of $R(t)$ in (6.12) we get $\phi := \tilde{\phi}$. We can thus substitute $\hat{q}$ and $\hat{\tilde{\eta}}$ in expression (6.4),(6.5) in lieu of q and $\tilde{\eta}(t,z)$ respectively. With the choice of control $\hat{u}$ given by (6.8) and boundedness condition on the market parameters r, μ and σ, the conditions in Theorem 4.1 are satisfied and hence $\hat{u}(t)$ is an optimal control process and the explicit representation of $\hat{p}$ is given by

$$\hat{p}(t) = (X(t))^{\gamma-1} e^{E[\exp(\int_t^T \gamma r(s,\theta(s)) - \bar{m}^2(s,\theta(s)) + \frac{(2-\gamma)}{(1-\gamma)} \frac{\bar{m}^2(s,\theta(s))}{2\sigma^2(s,\theta(s))} ds | \theta(t-)=i, Y(t-)=y)]}.$$

Quadratic loss minimization We now provide an example related to quadratic loss minimization where the portfolio wealth process is given by

$$dX^h(t) = \left(r(t,\theta(t))X^h(t) + h(t)\sigma(t,\theta(t))\bar{m}(t,\theta(t)) - h(t) \int_\Gamma g(t, X^h(t), \theta(t), \gamma)\pi(d\gamma) \right) dt$$
$$+ \ h(t)\sigma(t,\theta(t))dW(t) + h(t) \int_\Gamma g(t, X^h(t), \theta(t), \gamma)\tilde{N}(dt, d\gamma),$$
$$X^h(0) = x_0 \ a.s. \tag{6.15}$$

where the market price of risk is defined as $\bar{m}(t,i,y) = \sigma^{-1}(t,i)(b(t,i) - r(t,i))$. As like earlier example , we have that $\bar{m}(t,i) \geq 0$ and that the variables $r(t,i), b(t,i), \sigma(t,i)$, $\sigma^{-1}(t,i)$ and $g(t,x,i,\gamma)$ for each $i \in \mathcal{X}$ are measurable and uniformly bounded in $t \in [0,T]$. We assume that $g(t,x,i,\gamma) > -1$ for each $i \in \mathcal{X}$ and for a.a. t, x, γ. This insures that $X^h(t) > 0$ for each t. We further assume following conditions for each $i \in \mathcal{X}$

1. $E[\int_0^T h^2(t)dt] < \infty$.

2. $E[\int_0^T |r(t,i)X(t) + h(t)\sigma(t,i)\bar{m}(t,i)|dt + \int_0^T h^2(t)\sigma^2(t,i)dt + \int_0^T h^2(t)g^2(t,X(t),i,\gamma)dt] < \infty$.

3. $t \to \int_{\mathbb{R}} h^2(t)g^2(t,x,i,\gamma)\pi(d\gamma)$ is bounded.

4. the SDE for X has a unique strong solution.

The portfolio process $h(\cdot)$ satisfying the above four conditions is said to be admissible and belongs to $\mathcal{A}(T)$ (say). We consider the problem of finding an admissible portfolio process $h \in \mathcal{A}(T)$ such that

$$\inf_{h \in \mathcal{A}(T)} E[(X^h(T) - d)^2],$$

over all $h \in \mathcal{A}(T)$. Set the control process $u(t) \triangleq h(t)$ and $X(t) \triangleq X^h(t)$. For this example the Hamiltonian (3.4) becomes

$$
\mathcal{H}(t,x,h,i,y,p,q,\eta) = \left[r(t,i)x + u\sigma(t,i)\bar{m}(t,i) - u\int_\Gamma g(t,x,i,\gamma)\pi(d\gamma)\right]p + u\sigma(t,i)q
$$
$$
+ \left(u\int_\Gamma g(t,x,i,\gamma)\pi(d\gamma)\right)\eta, \tag{6.16}
$$

and the adjoint equations are for all time $t \in [0,T)$,

$$
dp(t) = -r(t,\theta(t-))p(t)dt + q(t)dW(t) + \int_\Gamma \eta(t,\gamma)\tilde{N}(dt,d\gamma) + \int_{\mathbb{R}} \tilde{\eta}(t,z)\tilde{N}_1(dt,dz),
$$
$$
p(T) = -2X(T) + 2d \quad a.s. \tag{6.17}
$$

We seek to determine $p(t), q(t), \eta(t,\gamma)$ and $\tilde{\eta}(t,z)$ in (6.17). Going by (6.17) we assume that ,

$$p(t) = \phi(t,\theta(t),Y(t))X(t) + \psi(t,\theta(t),Y(t)). \tag{6.18}$$

with the terminal boundary conditions being

$$\phi(T,i,y) = -2 \qquad \psi(T,i,y) = 2d \quad \forall\ i\ \in\ \mathcal{X}. \tag{6.19}$$

For the sake of convenience we again rewrite the following Ito's formula for a function $f(t,\theta(t),y(t)) \in \mathcal{C}^{1,2,1}$ given as

$$
df(t,\theta(t),Y(t)) = \left(\frac{\partial f(t,\theta(t),Y(t))}{\partial t} + \frac{(f^h(y/i))}{(1-F^h(y/i))}\sum_{j\neq i,j=1}^M p_{\theta(t-)=i,j}[f(t,j,0) - f(t,\theta(t-),y)]\right.
$$
$$
+ \left.\frac{\partial f(t,\theta(t),Y(t))}{\partial y}\right)dt
$$
$$
+ \int_{\mathbb{R}}[f(t,\theta(t-) + \bar{h}(\theta(t-),Y(t-),z),Y(t-) - \bar{g}(\theta(t-),Y(t-),z)) - f(t,\theta(t-),Y(t-))]\tilde{N}_1(dt,dz).
$$
$$\tag{6.20}$$

21

We apply the Ito's product rule to (6.18) to obtain

$$
\begin{aligned}
dp(t) &= X(t-)d\phi(t,\theta(t-),Y(t)) + \phi(t,\theta(t-),Y(t))dX(t) + d\phi(t,\theta(t-),Y(t))dX(t) + d\psi(t)\\
&= \sum_{i=1}^{M} 1_{\theta_{t-}=i}\Bigg\{ X(t-)\Big(\phi(t,\theta(t-),y)r(t,\theta(t-)) + \phi_t(t,\theta(t-),Y(t)) + \phi_y(t,\theta(t-),Y(t))\\
&\quad + \sum_{i=1,j\neq i}^{M} p_{ij}\frac{f^h(y/i)}{1-F^h(y/i)}(\phi(t,j,0)-\phi(t,\theta(t-),Y(t)))\Big) + u(t)\phi(t,\theta(t-),Y(t))\sigma(t,\theta(t-))\bar{m}(t,\theta(t-))\\
&\quad - u(t)\phi(t,\theta(t-),Y(t))\int_{\Gamma} g(t,X(t),\theta(t-),\gamma)\pi(d\gamma) + \psi_t(t,\theta(t-),Y(t)) + \psi_y(t,\theta(t-),Y(t))\\
&\quad + \sum_{i=1,i\neq j}^{M} p_{ij}\frac{f^h(y/i)}{1-F^h(y/i)}[\psi(t,j,0)-\psi(t,\theta(t-)=i,Y(t))]\Bigg\}dt\\
&\quad + u(t)\phi(t,\theta(t-),Y(t))\sigma(t,\theta(t-))dW(t) + u(t)\phi(t,\theta(t-),Y(t-))\int_{\Gamma} g(t,X(t-),\theta(t-),\gamma)\tilde{N}(dt,d\gamma)\\
&\quad + \int_{\mathbb{R}}\Big[X(t-)(\phi(t,\theta(t-)+\bar{h}(\theta(t-),Y(t-),z),Y(t-)-\bar{g}(\theta(t-),Y(t-),z)) - \phi(t,\theta(t-),Y(t-)))\\
&\quad + \psi(t,\theta(t-)+\bar{h}(\theta(t-),Y(t-),z),Y(t-)-\bar{g}(\theta(t-),Y(t-),z)) - \psi(t,\theta(t-),Y(t-))\Big]\tilde{N}_1(dt,dz).
\end{aligned}
$$

$$\tag{6.21}$$

Comparing coefficients with (6.17) we obtain three equations given as

$$
\begin{aligned}
&- r(t,\theta(t-))p(t-)\\
&= \sum_{i=1}^{M} 1_{\{\theta_{t-}=i,Y(t-)=y\}}\Bigg\{ X(t-)\Big(\phi(t,\theta(t-),Y(t))r(t,\theta(t-)) + \phi_t(t,\theta(t-),Y(t)) + \phi_y(t,\theta(t-),Y(t))\\
&\quad + \sum_{i=1,j\neq i}^{M} p_{ij}\frac{f^h(y/i)}{1-F^h(y/i)}(\phi(t,j,0)-\phi(t,\theta(t-),Y(t)))\Big) + u(t)\phi(t,\theta(t-),Y(t))\sigma(t,\theta(t-))\bar{m}(t,\theta(t-))\\
&\quad - u(t)\phi(t,\theta(t-),Y(t))\int_{\Gamma} g(t,x,\theta(t-),\gamma)\pi(d\gamma) + \psi_t(t,\theta(t-)),Y(t) + \psi_y(t,\theta(t-),Y(t))\\
&\quad + \sum_{i\neq j}^{M} p_{ij}\frac{f^h(y/i)}{1-F^h(y/i)}[\psi(t,j,0)-\psi(t,\theta(t-),Y(t))]\Bigg\}.
\end{aligned}
$$

$$\tag{6.22}$$

$$q(t) = u(t)\phi(t,\theta(t-),Y(t-))\sigma(t,\theta(t-)).\tag{6.23}$$

$$\eta(t,\gamma) = u(t)\phi(t,\theta(t-),Y(t-))g(t,X(t-),\theta(t-),\gamma).\tag{6.24}$$

$$
\begin{aligned}
\tilde{\eta}(t,z) &= X(t-)(\phi(t,\theta(t-)+\bar{h}(\theta(t-),Y(t-),z),Y(t-)-\bar{g}(\theta(t-),Y(t-),z)) - \phi(t,\theta(t-),Y(t-)))\\
&\quad + \psi(t,\theta(t-)+\bar{h}(\theta(t-),Y(t-),z),Y(t-)-\bar{g}(\theta(t-),Y(t-),z)) - \psi(t,\theta(t-),Y(t-)).
\end{aligned}
$$

$$\tag{6.25}$$

22

Let $\hat{u} \in \mathcal{A}(T)$ be a candidate optimal control corresponding to the wealth process $\hat{X}(T)$ and the adjoint triplet $(\hat{p}, \hat{q}, \hat{\eta}, \hat{\tilde{\eta}})$. Then from the Hamiltonian (3.4) for all $u \in \mathcal{A}(T)$ we have

$$
\begin{aligned}
\mathcal{H}(t, \hat{X}(t), u, \theta(t), \hat{p}(t), \hat{q}(t), \hat{\eta}(t)) &= \Big(r(t, \theta(t))\hat{X}(t) + u\sigma(t, \theta(t))\bar{m}(t, \theta(t)) \\
&\quad - u \int_\Gamma g(t, \hat{X}(t-), \theta(t-), \gamma)\pi d(\gamma) \Big)\hat{p}(t) \\
&\quad + u\sigma(t, \theta(t))\hat{q}(t) + \Big(u \int_\Gamma g(t, \hat{X}(t-), \theta(t-), \gamma)\pi(d\gamma) \Big)\hat{\eta}(t, \gamma).
\end{aligned}
\tag{6.26}
$$

As this is a linear function of u, we guess that the coefficient of u vanishes at optimality, which results in the following equality

$$
\begin{aligned}
\hat{q}(t) &= \Big(-\bar{m}(t, \theta(t-)) + \frac{1}{\sigma(t, \theta(t-))} \int_\Gamma g(t, \hat{X}(t), \theta(t), \gamma)\pi(d\gamma) \Big)\hat{p}(t) \\
&\quad - \frac{1}{\sigma(t, \theta(t-))} \int_\Gamma (g'(t, \hat{X}(t), \theta(t), \gamma))\pi(d\gamma)\hat{\eta}(t, \gamma).
\end{aligned}
\tag{6.27}
$$

Also substituting (6.27) for $\hat{q}(t)$ in (6.23) and using (6.18) and(6.24) we get,

$$
\hat{u}(t) = \frac{\tilde{\Lambda}(t)}{\Lambda(t)}(\hat{X}(t) + \phi^{-1}(t, \theta(t-), y)\psi(t, \theta(t-), y)),
\tag{6.28}
$$

where

$$
\tilde{\Lambda}(t) = -\bar{m}(t, \theta(t-))\sigma(t, \theta(t-)) + \int_\Gamma g(t, X(t), \theta(t-), \gamma)\pi(d\gamma).
$$

$$
\Lambda(t) = \sigma^2(t, \theta(t-)) + \phi(t, \theta(t-), Y(t)) \int_\Gamma g'(t, X(t), \theta(t-), \gamma)g(t, X(t), \theta(t-), \gamma)\pi(d\gamma).
\tag{6.29}
$$

To find the optimal control it remains to find ϕ and ψ. To do so set $X(t) := \hat{X}(t), u(t) := \hat{u}(t)$ and $p(t) := \hat{p}(t)$ in (6.22) and then substitute for $\hat{p}(t)$ in (6.18) and $\hat{u}(t)$ from (6.28) . As this result is linear in $\hat{X}(t)$ we compare the coefficient on both side of the resulting equation to get following two equations namely,

$$
\begin{aligned}
0 &= 2r\phi(t, i, Y(t)) + \phi_t(t, i, Y(t)) + \phi_y(t, i, Y(t)) + \sum_{i \neq j, i=1}^{M} p_{ij}\frac{f^h(y/i)}{1 - F^h(y/i)}(\phi(t, j, 0) - \phi(t, i, Y(t))) \\
&\quad + \frac{\tilde{\Lambda}(t)}{\Lambda(t)}\sigma(t, i)\bar{m}(t, i)\phi(t, i, Y(t)) - \frac{\tilde{\Lambda}(t)}{\Lambda(t)}\phi(t, i, Y(t)) \int_\Gamma g(t, X(t), i, \gamma)\pi(d\gamma).
\end{aligned}
\tag{6.30}
$$

$$
\begin{aligned}
0 &= r\psi(t, i, Y(t)) + \psi_t(t, i, Y(t)) + \psi_y(t, i, Y(t)) + \sum_{i \neq j, i=1}^{M} p_{ij}\frac{f^h(y/i)}{1 - F^h(y/i)}(\psi(t, j, 0) - \psi(t, i, Y(t))) \\
&\quad + \frac{\tilde{\Lambda}(t)}{\Lambda(t)}\sigma(t, i)\bar{m}(t, i)\psi(t, i, Y(t)) - \frac{\tilde{\Lambda}(t)}{\Lambda(t)}\psi(t, i, y) \int_\Gamma g(t, X(t), i, \gamma)\pi d(\gamma).
\end{aligned}
\tag{6.31}
$$

23

with terminal boundary conditions given by (6.19). Consider the following process

$$\tilde{\phi}(t,i,y) = -2E\left[\exp\left\{\int_t^T \left(2r(s,\theta(s-)) + \frac{\tilde{\Lambda}(s)}{\Lambda(s)}\sigma(s,\theta(s-))\bar{m}(s,\theta(s-))\right.\right.\right.$$
$$\left.\left.\left. -\frac{\tilde{\Lambda}(s)}{\Lambda(s)}\int_\Gamma g(s,X(s),\theta(s-),\gamma)\pi(d\gamma)\right)ds\right\}|(\theta(s-)=i,Y(t)=y)\right].$$
(6.32)

$$\tilde{\psi}(t,i,y) = 2dE\left[\exp\left\{\int_t^T \left(r(\theta(s-),s) + \frac{\tilde{\Lambda}(s)}{\Lambda(s)}\sigma(s,\theta(s-))\bar{m}(s,\theta(s-))\right.\right.\right.$$
$$\left.\left.\left. -\frac{\tilde{\Lambda}(s)}{\Lambda(s)}\int_\Gamma g(s,X(s),\theta(s-),\gamma)\pi(d\gamma)\right)ds\right\}\bigg|(\theta(s-)=i,Y(s)=y)\right].$$
(6.33)

We aim to show that $\phi = \tilde{\phi}$ and $\psi = \tilde{\psi}$. We define the following martingales:

$$R(t) = E\left[\exp\left\{\int_0^T \left(2r(s,\theta(s-)) + \frac{\tilde{\Lambda}(s)}{\Lambda(s)}\sigma(s,\theta(s-))\bar{m}(s,\theta(s-)) - \frac{\tilde{\Lambda}(s)}{\Lambda(s)}\int_\Gamma g(s,X(s),\theta(s-),\gamma)\pi(d\gamma)\right)ds\right\}|\mathcal{F}_t^{\theta,y}\right],$$
(6.34)

$$S(t) = E\left[\exp\left\{\int_0^T \left(r(s,\theta(s-)) + \frac{\tilde{\Lambda}(s)}{\Lambda(s)}\sigma(s,\theta(s-))\bar{m}(s,\theta(s-)) - \frac{\tilde{\Lambda}(s)}{\Lambda(s)}\int_\Gamma g(s,X(s),\theta(s-),\gamma)\pi(d\gamma)\right)ds\right\}|\mathcal{F}_t^{\theta,y}\right],$$
(6.35)

where $\mathcal{F}_t^{\theta,y}$ is defined as usual. We follow steps similar to that as seen in Example 1 and conclude that $\phi = \tilde{\phi}$ and $\psi = \tilde{\psi}$ by using joint-Markov property of $(\theta(t),Y(t))$, to obtain the following expression for the control $\hat{u}(t)$ given as

$$\hat{u}(t) = \frac{\tilde{\Lambda}(t)}{\Lambda(t)}\left(\hat{X}(t) - \frac{dE\left[\exp\left\{\int_t^T (r(s,\theta(s-))+\frac{\tilde{\Lambda}(s)}{\Lambda(s)}\sigma(s,\theta(s-))\bar{m}(s,\theta(s-))-\frac{\tilde{\Lambda}(s)}{\Lambda(s)}\int_\Gamma g(s,X(s),\theta(s-),\gamma)\pi(d\gamma))ds\right\}|(\theta(t-)=i,Y(t)=y)\right]}{E\left[\exp\left\{\int_t^T (2r(s,\theta(s-))+\frac{\tilde{\Lambda}(s)}{\Lambda(s)}\sigma(s,\theta(s-))\bar{m}(s,\theta(s-))-\frac{\tilde{\Lambda}(s)}{\Lambda(s)}\int_\Gamma g(s,X(s),\theta(s-),\gamma)\pi(d\gamma))ds\right\}|(\theta(t)=i,Y(t)=y)\right]}\right).$$

For the choice of the control parameter and the boundedness conditions on the market parameters r, b, σ and g, the conditions of Theorem 4.1 are satisfied and hence $\hat{u}$ is the optimal control process.

References

[1] Cadenillas, A., and Karatzas, I. 1995. The stochastic maximum principle for linear, convex optimal control with random coefficients. *SIAM J. Control Optim.* 33(2): 590-624.

[2] Cohen, S.N., Elliott, R.J. 2010. Comparisons for backward stochastic differential equations on Markov chains and related no-arbitrage conditions. *Ann. Appl. Probab.* 20(1): 267-311.

[3] Deshpande, A., and Ghosh M.K. 2008. Risk minimizing option pricing in a Markov modulated market. *Stochastic Anal.Appl.* 26(1): 313–324.

[4] Donnelly, C. 2011. Sufficient Stochastic Maximum Principle in a Regime-Switching Diffusion Model. *Appl Math Optim.* 64(2):155-169.

[5] Framstad N.C., Oksendal,B., and Sulem, A. 2004. Sufficient stochastic maximum principle for the optimal control of jump diffusions and applications to finance. *J. Optim. Theory Appl.*121(1): 77-98.

[6] Ghosh, M.K., and Goswami, A. 2009. Risk minimizing option pricing in a semi-Markov modulated market. *SIAM J. Control Optim.* 48(3): 1519–1541.

[7] Gikhman, I.I., and Skorokhod, A.V. 1983. *The Theory of Stochastic Processes II*, Springer, Berlin.

[8] Ikeda, N., and Watanabe, S. 1981. *Stochastic differential equations and diffusion processes*, North-Holland, Amsterdam.

[9] Karatzas, I., and Shreve, S. 1998. *Methods of Mathematical Finance*, Springer, New York.

[10] Oksendal, B., and Sulem,A. 2007. *Applied Stochastic Control of Jump Diffusions*, Second Edition, Universitext, Springer Verlag, Berlin, Heidelberg.

[11] Zhang,X., Elliott,R.J., and Siu, T.K. 2012. A Stochatic maximum principle for a Markov regime switching jump-diffusion model and its application to finance. *SIAM J. Control Optim.* 50 (2): 964-990.

[12] Whittle,P. 1980. *Risk sensitive Optimal control*, John Wiley and Sons.

[13] Zhou, X.Y., and Yin G. 2003. Markowitzs mean-variance portfolio selection with regime switching: A continuous-time model. *SIAM J. Control Optim.* 42(4): 1466-1482.