Approximation of Optimal Ergodic Dividend and Reinsurance Strategies Using Controlled Markov Chains

Zhuo Jin,* Hailiang Yang,† G. Yin‡

November 13, 2017

Abstract

This work develops a numerical method for finding optimal ergodic (long-run average) dividend and reinsurance strategies in a regime-switching model. The surplus process is modelled by a Markov switching process subject to liability constraints. Using dynamic programming principle, the optimal long term average dividend payment is a solution of coupled system of Hamilton-Jacobi-Bellman equations. Under suitable conditions, the optimal value of long-term average dividend payment can be represented by using an invariant measure. However, due to the regime-switching, approximating the invariant measure is very difficult. Our goal is to design a numerical algorithm to approximate the optimal ergodic dividend payment strategy directly. We use Markov chain approximation techniques to construct a discrete-time controlled Markov chain for the approximation. Convergence of the approximation algorithms is proved. Examples are presented to illustrate the applicability of the numerical methods.

Key Words. Stochastic control, ergodic dividend payment strategy, reinsurance, invariant measure, regime-switching model, Markov chain approximation.

*Centre for Actuarial Studies, Department of Economics, The University of Melbourne, VIC 3010, Australia, zjin@unimelb.edu.au.

†Department of Statistics and Actuarial Science, The University of Hong Kong, Hong Kong, hlyang@hku.hk.

‡Department of Mathematics, Wayne State University, Detroit, Michigan 48202, gyin@math.wayne.edu.

1 Introduction

Managing the surplus and designing dividend payment policies have long been an important issue in finance and actuarial sciences. The dividend payment plan released in the financial report for public companies represents an important signal about a firm's future growth opportunities and profitability. The decision of the dividend payment is so important for a public company's financial strength because the company's share price is very sensitive to the information of dividend plans, and dividend payment strategies also influence the investment and financing decisions of firms. For insurance companies, because of the nature of their products, insurers tend to accumulate relatively large amounts of cash, cash equivalents, and investments in order to pay future claims and avoid insolvency. The payment of dividends to shareholders may reduce an insurer's ability to survive adverse investment and underwriting experience. However, due to the undergone pressure of managing balance sheets and distributing the surplus for public insurance companies, one natural objective for insurers is to optimize the management of surplus and sustained stream of dividend payments. A practitioner will manage the reserve and dividend payment against various financial risks so that the company can satisfy its minimum capital requirement in the long run.

Since the introduction of the optimal dividend payment model proposed by De Finetti (1957), there have been increasing efforts on using advanced methods from the toolbox of stochastic control theory to study the optimal dividend policy. The majority of research are conducted in finding optimal dividend payment policies to maximize the present value of the cumulative dividend payment in targeted time horizon. Regular controls, singular controls, and impulse controls are involved in various scenarios. Guo et al. (2004) studies the dividend and Risk control with a diffusion where the drift is quadratic in the risk control variable. He and Liang (2008) studies the mixed control of dividend, proportional reinsurance and financing for the model. Løkka and Zervos (2008) solves the optimal dividend and issuance of equity policies in the presence bankruptcy risk. Meng and Siu (2011) and Wei et al. (2014) study the combined control of dividend, financing and risk for the Brownian motion model. Alvarez and Lempa (2008) and Bai and Paulsen (2012) study the impulse dividend control problem for a rather general linear diffusion model in which some growth and smoothness conditions are imposed on. Avram et al. (2007) addresses the dividend and reinvestment control in a spectrally negative Lévy process. Azcue and Muler (2010) analyzes the problem of the maximization of total discounted dividend payment for an insurance company. Loeffen and Renaud (2010) investigates the optimal dividend control with affine penalty at ruin for a spectrally negative Lévy process using singular control. Jin et al. (2015) considers the credit risk and derives the optimal debt ratio and dividend payment strategies for an insurance company.

In previous work that mainly follows the classical Cramér-Lundberg risk model with dividends payments, when total discounted dividend payment is maximized, the company will almost surely be financially ruined. In Jin et al. (2015), the authors propose an asset and liability model with liability constraint. The insurance company manages the surplus and designs the dividend payment strategies taking into account the liability capacity. Then, the insurance company will be in the absence of insolvency. On the other hand, since insurance companies generally have assets and liabilities with long maturity, in particular, for life insurance companies, it is very important for insurance companies to consider long-term objectives and build the dividend payment strategies with long time horizon in mind. Long-term objectives are widely studied in investment and risk management in a variety of cases. Bielecki and Pliska (1999) analyzes the dynamic asset management in an infinite time horizon. Fleming and Sheu (2000) studies the optimal investment strategies with a long-term objective. Pham (2003) proposes a large deviation approach for optimal long term investment. See also Bielecki and Pliska (2003) and Fleming and McEneaney (1995) for related works. In this work, we extend the asset and liability model in Jin et al. (2015) and set a new objective function to consider the long-term impact of the dividend payment strategies, which is applicable when financial ruin is completely avoided. Instead of adopting the discounted present value, we aim to maximize the average dividend payment in the long term.

Further, people have recently realized that stochastic hybrid surplus models have advantages to capture discrete movements (such as random environment, market trends, interest rates, business cycles, etc.). The hybrid system investigates the coexistence of continuous dynamics and discrete events in the systems. To reflect the hybrid feature, one of the recent trends is to use a finite state Markov process to describe the transitions among different regimes. The Markov-modulated switching systems are therefore known as regime-switching systems. Thus the formulation of regime-switching models is a more general and versatile framework to describe the complicated financial markets and their inherent uncertainty and randomness. In Wei et al. (2010), the optimal dividend and proportional reinsurance strategy under utility criteria are studied for the regime-switching compound Poisson model. Sotomayor and Cadenillas (2011) studies the optimal dividend problem in the regime-switching model when the dividend rates are bounded, unbounded, and when there are fixed costs and taxes corresponding to the dividend payments. Zhu (2014a) studied the dividend optimization for a regime-switching diffusion model with restricted dividend rates. See also Zhu and Chen (2013), Zhu and Yang (2015), and Jin et al. (2011). A comprehensive study of switching diffusions with "state-dependent" switching is in Yin and Zhu (2010).

To find the optimal strategies, one usually solves a so-called Hamilton-Jacobi-Bellman (HJB) equation. However, because of the regime-switching jump diffusion, the HJB equation is in fact a coupled system of nonlinear HJB equations. To represent the maximal average dividend payment in the long term, the unique invariant measure is constructed. Due to the complexity of the Markov switches, the explicit formula of the long-term average is virtually impossible to obtain. A viable alternative is to employ numerical approximations. We adopt the Markov chain approximation methodology developed in Kushner and Dupuis (2001) to solve for the optimal performance function and the corresponding dividend payment strategies. A numerical algorithm for approximating optimal investment and dividend payment policies with capital injections under regime-switching diffusion models is developed in Jin et al. (2013). In this work, we carry out a convergence analysis for our formulation using weak convergence methods and relaxed control formulation of numerical schemes in the setting of regime switching, in which case one needs to deal with a system of HJB equations with reflecting boundaries. Comparing with the work in Jin et al. (2013), we choose a different performance function, which performs in an infinite time horizon and require the ergodicity of the diffusion process. The long-run average objective function adds much difficulties to design the numerical schemes and to conduct the the convergence analysis of the algorithm. The numerical implementation can be done using either value iterations or policy iterations. It is worth to mention that the Markov chain approximation method requires little regularity of the objective function and/or analytic properties of the associated systems of HJB equations.

The rest of the paper is organized as follows. A general formulation of surplus process, performance functions, and assumptions are presented in Section 2. The existence and uniqueness of the invariant measure are presented. The dynamic programming equation is derived in Section 4. Section 5 deals with the numerical algorithm of Markov chain approximation method. Section 6 deals with the convergence of the approximation scheme. The technique of "rescaling time" is introduced and the convergence theorems are proved. Numerical examples are presented in Section 7 to illustrate the performance of the approximation method together with some further remarks.

2 Formulation

For an insurance company, the surplus process $X(t)$ is described as the difference between the asset value $A(t)$ and liabilities $L(t)$. That is,

$$X(t) = A(t) - L(t). \tag{2.1}$$

To delineate the random environment and other random factors, we use a continuous-time irreducible Markov chain $\alpha(t)$ taking values in the finite space $\mathcal{M} = \{1, \ldots, m\}$. The market states are represented by the Markov chain $\alpha(t)$, and they undergo a Markov regime switching. Let the continuous-time Markov chain $\alpha(t)$ be generated by $Q = (q_{ij}) \in \mathbb{R}^{m\times m}$. That is,

$$\mathbb{P}\{\alpha(t+\delta) = j | \alpha(t) = i, \alpha(s), s \le t\} = \begin{cases} q_{ij}\delta + o(\delta), & \text{if } j \neq i, \\ 1 + q_{ii}\delta + o(\delta), & \text{if } j = i, \end{cases}$$

where $q_{ij} \ge 0$ for $i, j = 1, 2, \ldots, m$ with $j \neq i$ and $q_{ii} = -\sum_{j\neq i} q_{ij} < 0$ for each $i = 1, 2, \ldots, m$. We further assume that Q is irreducible.

When the insurer incurs a liability at time t, he receives a premium for the amount insured. The collected premium will increase assets and surplus at time t. Denote by $\beta(\alpha(t))$ the premium rate, where for each $i \in \mathcal{M}$, $\beta(i)$ represents the cost of protection per dollar of insurance liabilities. The asset value increases from the insurance sales during the time period $[t, t + dt]$ is denoted as $\beta(\alpha(t))L(t)dt$.

To protect insurance companies against the impact of claim volatilities, reinsurance is a standard tool with the goal of reducing and eliminating risks. The primary insurance carrier pays the reinsurance company a certain part of the premiums. In return, the reinsurance company is obliged to share the risk of large claims. We assume that proportional reinsurance is adopted by the primary insurance company in our model. Within this scheme, the reinsurance company covers a fixed percentage of losses. Let λ be an exogenous retention level for the reinsurance policy. Note that $\lambda \in [0, 1]$. Denote by $h(\lambda)$ be reinsurance charge rate (the cost of reinsurance protection per dollar of reinsured liabilities) for hedging the adverse claims due to the downside risk of the securities' values. From a practical view of point, the cost of reinsurance protection per dollar of reinsured liabilities should be nonnegative and less than 1. Thus, we assume that $h(\lambda)$ is bounded and $h(\lambda) \in [0, 1]$. Hence, the reinsurance charge during the time period $[t, t+dt]$ is $h(\lambda)L(t)dt$, and only $\lambda L(t)dt$ will be covered by the primary insurance company.

At this premium rate α and reinsurance retention level λ, there is an elastic demand for insurance contract and the insurer decides how much insurance $L(t)$ to offer at that premium rate and reinsurance retention level. Let $\pi(t) = L(t)/X(t)$ be the debt ratio of the insurance company. Then, the leverage, which is described as the ratio between asset values and surplus, can be written as $A(t)/X(t) = 1 + \pi(t)$. To avoid the insurance liabilities being too large, the insurers will decide the optimal liabilities to manage the sale of insurance policies.

We assume that the asset value $A(t)$ in the financial market follows a geometric Brownian motion process

$$\frac{dA(t)}{A(t)} = \mu(\alpha(t))dt + \sigma(\alpha(t))dw(t), \tag{2.2}$$

where for each $i \in \mathcal{M}$, $\mu(i)$ is the return rate of the asset and $\sigma(i)$ is the corresponding volatility and $w(t)$ is a standard Brownian motion. Hence, combining (2.1)–(2.2), the surplus process in the absence of claims and dividend payment can be denoted by $\widetilde{X}(t)$ such that

$$d\widetilde{X}(t) = (\beta(\alpha(t)) - h(\lambda))L(t)dt + A(t)(\mu(\alpha(t))dt + \sigma(\alpha(t))dw(t)). \tag{2.3}$$

We further consider the claims, which are against insurer's liabilities incurred earlier. Denoted by $R(t)$ the future claims up to time t. Then we assume that the claims are proportional to the amount of insurance liabilities $L(t)$. Hence, the accumulated claims up to time T is denoted as

$$R(T) = \int_0^T c(t)L(t)dt, \tag{2.4}$$

where $c(t)$ can be considered as a claim rate against liabilities.

Practically, the claim rate $c(t)$ is risky and is not predictable. The claim rates of different types of insurance products are very different and volatile in different types of markets. For example, the CDS, an insurance contract to protect against credit events, is affected by a series of economic factors such as credit ratings of banks and insurance companies, government regulation, and demand of CDOs in the market, etc. In addition, it is largely influenced by the randomness of economic environment that are described as random shocks. In Jin et al. (2015), the claim rate $c(t)$ is formulated as a diffusion process to describe its randomness. However, one of the main drawbacks of the diffusion process in the work is that the claim rate can be negative, which is difficult to calibrate with market data. To guarantee the positivity of the claim, we assume that the claim rate follows a continuous-time Markov process, taking values in a set of positive values. That is, the claim rate depends on $\alpha(t)$, so $c(\alpha(t))$ in lieu of $c(t)$ is used. Hence, the accumulated claims follows

$$R(T) = \int_0^T c(\alpha(t))L(t)dt \tag{2.5}$$

A dividend strategy $D(\cdot)$ is an $\mathcal{F}_t$-adapted process $\{D(t) : t \geq 0\}$ corresponding to the accumulated amount of dividends paid up to time t such that $D(t)$ is a nonnegative and nondecreasing stochastic process that is right continuous and have left limits with $D(0^-) = 0$. In this paper, we consider the optimal dividend strategy where the dividend payments are proportional to the surplus with a dividend payment rate $u(t)$. Denote $U = [0, 1]$. As a result, we write $D(t)$ as

$$dD(t) = u(t)X(t)dt, \tag{2.6}$$

where $u(t)$ is an $\mathcal{F}_t$-adapted process and $0 \leq u(t) \leq 1$. Thus, taking into consider the impact of reinsurance, the insurer's surplus process in the presence of claims and dividend payments is given by

$$dX(t) = d\widetilde{X}(t) - \lambda dR(t) - dD(t). \tag{2.7}$$

Together with the initial condition, (2.7) follows

$$\begin{cases} dX(t) = [(\beta(\alpha(t)) - h(\lambda) - \lambda c(\alpha(t)))L(t) + \mu(\alpha(t))A(t) - u(t)X(t)]dt + A(t)\sigma(\alpha(t))dw(t), \\ X(0) = x \geq 0 \end{cases} \tag{2.8}$$

for all $t < \tau$ and we impose $X(t) = 0$ for all $t > \tau$, where $\tau = \inf\{t \geq 0 : X(t) < 0\}$ represents the time of financial ruin. Suppose the optimal payout strategy is applied subsequently.

Recall that $\pi(t)$ represents the debt ratio, (2.8) can be written as

$$\begin{cases} \dfrac{dX(t)}{X(t)} = [\pi(t)(\beta(\alpha(t)) - h(\lambda) - \lambda c(\alpha(t)) + \mu(\alpha(t))) + \mu(\alpha(t)) - u(t)]dt + (\pi(t) + 1)\sigma(\alpha(t))dw(t), \\ X(0) = x. \end{cases} \tag{2.9}$$

For dividend payment rate, $u(t)$ is non-negative and subject to an upper bound. A strategy $u(\cdot)$ being progressively measurable with respect to $\{w(s), \alpha(s) : 0 \leq s \leq t\}$ is called an admissible strategy. Denote the collection of all admissible strategies or admissible controls by $\mathcal{U}$. Then the admissible strategy set $\mathcal{U}$ can be defined as

$$\mathcal{U} = \Big\{u \in \mathbb{R} : 0 \leq u \leq 1\Big\}. \tag{2.10}$$

A Borel measurable function $u(x, \alpha)$ is an admissible feedback strategy or feedback control if (2.9) has a unique solution.

The objective of the representative financial institute is to maximize the average dividend payment in the long term. For an arbitrary admissible feedback control $u(\cdot, \cdot)$, the performance function is average dividend payment in the long term given by

$$J(x, u, i) = \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}_{x,i}\Big[\int_0^T u(X(t), \alpha(t))X(t)dt\Big], \ \forall i \in \mathcal{M}, \tag{2.11}$$

where $\mathbb{E}_{x,i}$ denote the expectation conditioned on $X(0) = x$ and $\alpha(0) = i$, and let $\mathbb{P}_{x,i}$ denote the conditional probability on $X(0) = x$ and $\alpha(0) = i$.

Denote by $\gamma(u) = J(x, u, i)$. Define the optimal value as

$$\bar{\gamma} := \sup_{u\in\mathcal{U}} \gamma(u). \tag{2.12}$$

For an arbitrary $u \in \mathcal{U}$, $i = \alpha(t) \in \mathcal{M}$, and $V(\cdot, i) \in C^2(\mathbb{R})$, define an operator $\mathcal{L}^u$ by

$$\mathcal{L}^u V(x,i) = V_x(x,i)x(\pi(\beta(i)-h(\lambda)-\lambda c(i)+\mu(i))+\mu(i)-u)+\frac{1}{2}(u+1)^2\sigma^2(i)x^2V_{xx}(x,i)+QV(x,\cdot)(i) \tag{2.13}$$

where V_x and V_{xx} denote the first and second derivatives with respect to x, and

$$QV(x, \cdot)(i) = \sum_{j\neq i} q_{ij}(V(x, j) - V(x, i)).$$

If $\bar{\gamma}$ exists, by applying the dynamic programming principle (Fleming and Soner (2006)), there exists a sufficiently smooth function V that normally satisfies the following coupled system of HJB equations:

$$\bar{\gamma} = \max_{u\in\mathcal{U}}\{\mathcal{L}^u V(x, i) + ux\}, \quad \text{for each} \ \ i \in \mathcal{M}. \tag{2.14}$$

In view of of (2.9), the surplus is always nonnegative in the infinite time horizon. The insurance company will run the business with probability one in the long run. It is worthwhile to consider the ergodic control of dividend payment when the operation period $T \to \infty$. On the other hand, in reality, the surplus of an insurance company cannot reach infinity. Once the surplus is substantially high, the decision maker will undergo pressure from the shareholders to pay dividend. Hence, we need only choose B large enough and set the surplus in the finite interval $G = [0, B]$. To make $J(x, u, i)$ computationally feasible, we truncate x at some large value B.

3 Invariant Measure

To obtain the expected average dividend payment in the infinite time horizon, one approach is to replace the instantaneous measures with invariant measures. Note that the state of the process in our formulation has two component: one component is the diffusion process $X(t)$; the other component is the Markov regime switching process $\alpha(t)$. We denote by $Z(t) = (X(t), \alpha(t))$ the state of the process.

To proceed, we need the following assumption.

(A) $Z(t)$ is positive recurrent with respect to some bounded domain $E \times \{i\}$, where $E \subset G \subset \mathbb{R}$, i is fixed and $i \in \mathcal{M}$.

Lemma 3.1. *Assume* (A). *$Z(t)$ is positive recurrent with respect to $G \times \mathcal{M}$.*

Proof. The result is immediately obtained by applying Theorem 3.12 in Yin and Zhu (2010). □

We proceed to define a sequence of stopping times $\{\eta_k\}, k = 0, 1, 2, \ldots$ Let $\eta_0 = 0$, η_{2k+1} be the first time after η_{2k} when $Z(t)$ reaches the boundary $\partial E \times \{i\}$, and η_{2k+2} be the first time after η_{2k+1} when $Z(t)$ reaches the boundary $\partial G \times \{i\}$. Then, the sample path of $Z(t)$ can be divided to the cycles as

$$[\eta_0, \eta_2), [\eta_2, \eta_4), \ldots, [\eta_{2k}, \eta_{2k+2}), \ldots \tag{3.1}$$

By Lemma 3.1, $Z(t)$ is positive recurrent with respect to $G \times \mathcal{M}$. Hence the stopping times $\{\eta_k\}, k = 0, 1, 2, \ldots$ are finite almost surely. Without loss of generality, we assume $x = 0$. It follows that the sequence $Z_n = (X_n, i) = Z(\eta_{2n}), n = 0, 1, \ldots,$ is a Markov chain on $\partial G \times \{i\}$. Denote by $\mathcal{B}(\partial G)$ the collection of Borel measurable sets on ∂G. Starting from (x, i), $Z(t)$ may jump many times before it reaches the set (H, i) where $H \in \mathcal{B}(\partial G)$. The one-step transition probability of the Markov chain Z_n is defined as

$$\tilde{p}^{(1)}(x, H) = \mathbb{P}(Z_1 \in (H \times \{i\}) | Z_0 = (x, i)). \tag{3.2}$$

Analogously, the $n-$step transition probability of the Markov chain Z_n is denoted by $\tilde{p}^{(n)}(x, H)$. Now we will construct the stationary distribution of $Z(t)$.

Theorem 3.2. *The positive recurrent process $Z(t)$ has a unique stationary distribution $\nu(\cdot, \cdot)$. Let $\theta(\cdot, \cdot)$ be the stationary density associated with the stationary distribution. Then for any $(x, i) \in G \times \mathcal{M}$,*

$$\mathbb{P}_{x,i}\left(\lim_{T\to\infty} \frac{1}{T} \int_0^T u(X(t), \alpha(t))X(t)dt = \bar{\gamma}\right) = 1, \tag{3.3}$$

where

$$\bar{\gamma} = \sum_{i=1}^{m} \int_{\mathbb{R}} u(x, i)x\theta(x, i)dx. \tag{3.4}$$

Proof. In view of Lemma 4.1 in Yin and Zhu (2010), Z_n has a unique stationary distribution $\phi(\cdot)$. For any $H \in \mathcal{B}(\mathbb{R})$, $\phi(H) = \lim_{n\to\infty} \tilde{p}^{(n)}(x, H)$. Recall that the cycles are defined in (3.1). Denote by $\tau^{H\times\{i\}}$ the time spent by the path of $Z(t)$ in the set $(H \times \{i\})$ during the first cycle. Set

$$\tilde{\nu}(H, i) := \int_{\partial G} \phi(dx)\mathbb{E}_x \tau^{H\times\{i\}}. \tag{3.5}$$

Using Theorem 4.3 in Yin and Zhu (2010), we have

$$\begin{aligned}
&\sum_{i=1}^{m} \int_{\mathbb{R}} u(x, i)x\tilde{\nu}(dx, i) \\
&= \int_{\partial G} \phi(dx)\mathbb{E}_x \int_0^{\eta_2} u(X(t), \alpha(t))X(t)dt \\
&= \sum_{i=1}^{m} \int_{\mathbb{R}} \mathbb{E}_{x,i} u(X(t), \alpha(t))X(t)\tilde{\nu}(dx, i).
\end{aligned} \tag{3.6}$$

Hence, the desired stationary distribution is defined by the normalized measure as

$$\nu(H,i) = \frac{\tilde{\nu}(H,i)}{\sum_{j=1}^{m} \tilde{\nu}(\mathbb{R},j)}, \ \forall i \in \mathcal{M}. \tag{3.7}$$

Now we will prove (3.3). Regarding the stationary distribution, we know that starting from an arbitrary point (x,i) with arbitrary initial distribution is asymptotically equivalent to starting with the initial distribution being the stationary distribution. Then we will only need verify the case when the initial distribution is the stationary distribution of the Markov chain Z_n. That is, for any $H \in \mathbb{B}(\partial G)$,

$$\mathbb{P}\{(X(0),\alpha(0)) \in (H \times \{i\})\} = \phi(H).$$

Consider the sequence of random variables

$$\rho_n = \int_{\eta_{2n}}^{\eta_{2n+2}} u(X(t),\alpha(t))X(t)dt.$$

From (3.5) and (3.6), we have

$$\mathbb{E}\rho_n = \sum_{i=1}^{m} \int_{\mathbb{R}} u(x,i)x\tilde{\nu}(dx,i), \tag{3.8}$$

for all $n = 0, 1, 2, \ldots$ Let $\phi(T)$ denote the number of cycles completed up to time T. Then,

$$\phi(T) := \max\{n \in \mathbb{N} : \sum_{k=1}^{n} (\eta_{2k} - \eta_{2k-2} \le T)\}.$$

Hence, $\int_0^T u(X(t),\alpha(t))X(t)dt$ can be decomposed as

$$\int_0^T u(X(t),\alpha(t))X(t)dt = \sum_{n=0}^{\phi(T)} \rho_n + \int_{\eta_{2\phi(T)}}^{T} u(X(t),\alpha(t))X(t)dt.$$

Note that $u(\cdot)$ and $X(\cdot)$ both are nonnegative, we have

$$\sum_{n=0}^{\phi(T)} \rho_n \le \int_0^T u(X(t),\alpha(t))X(t)dt \le \sum_{n=0}^{\phi(T)+1} \rho_n.$$

Then,

$$\frac{1}{\phi(T)} \sum_{n=0}^{\phi(T)} \rho_n \le \frac{1}{\phi(T)} \int_0^T u(X(t),\alpha(t))X(t)dt \le \frac{1}{\phi(T)} \sum_{n=0}^{\phi(T)+1} \rho_n.$$

As $T \to \infty$, $\phi(T) \to \infty$. Combining with (3.8), we have

$$\frac{1}{\phi(T)} \int_0^T u(X(t),\alpha(t))X(t)dt \to \sum_{i=1}^{m} \int_{\mathbb{R}} u(x,i)x\tilde{\nu}(dx,i). \tag{3.9}$$

On the other hand, the law of large numbers implies

$$\mathbb{P}\left\{ \frac{1}{n} \sum_{k=0}^{n} \rho_k \to \sum_{i=1}^{m} \int_{\mathbb{R}} u(x,i)x\tilde{\nu}(dx,i), \text{ as } n \to \infty \right\} = 1. \tag{3.10}$$

Particulary, when $u(x,i)=1/x$, (3.10) implies

$$\mathbb{P}\left\{\frac{\eta_{2n+2}}{n}\to\sum_{i=1}^{m}\tilde{\nu}(dx,i),\ \text{as } n\to\infty\right\}=1. \tag{3.11}$$

Since $\eta_{2n}\le T\le\eta_{2n+2}$, and $\eta_{2n}/\eta_{2n+2}\to 1$ almost surely as $T\to\infty$, we have

$$\mathbb{P}\left\{\frac{T}{\phi(T)}\to\sum_{i=1}^{m}\tilde{\nu}(dx,i),\ \text{as } T\to\infty\right\}=1. \tag{3.12}$$

Now, using (3.9) and (3.12), we have as $T\to\infty$.

$$\begin{aligned}\frac{1}{T}\int_0^T u(t)X(t)dt &= \frac{\int_0^T u(t)X(t)dt}{\phi(T)}\frac{\phi(T)}{T}\\ &\to\sum_{i=1}^{m}\int_{\mathbb{R}}u(x,i)x\tilde{\nu}(dx,i)\times\frac{1}{\sum_{i=1}^{m}\tilde{\nu}(dx,i)}\\ &=\sum_{i=1}^{m}\int_{\mathbb{R}}u(x,i)x\nu(dx,i)\ \ \text{almost surely.}\end{aligned} \tag{3.13}$$

Hence,

$$\mathbb{P}\left(\lim_{T\to\infty}\frac{1}{T}\int_0^T u(X(t),\alpha(t))X(t)dt=\sum_{i=1}^{m}\int_{\mathbb{R}}u(x,i)x\nu(dx,i),\ \text{as } T\to\infty\right)=1. \tag{3.14}$$

Since (3.14) holds for any $(x,i)\in G\times\mathcal{M}$, then

$$\mathbb{P}_{x,i}\left(\lim_{T\to\infty}\frac{1}{T}\int_0^T u(X(t),\alpha(t))X(t)dt=\sum_{i=1}^{m}\int_{\mathbb{R}}u(x,i)x\nu(dx,i),\ \text{as } T\to\infty\right)=1. \tag{3.15}$$

Note that $\theta(\cdot,\cdot)$ is the stationary density associated with the stationary distribution $\nu(\cdot,\cdot)$, (3.15) can be written as

$$\mathbb{P}_{x,i}\left(\lim_{T\to\infty}\frac{1}{T}\int_0^T u(X(t),\alpha(t))X(t)dt=\sum_{i=1}^{m}\int_{\mathbb{R}}u(x,i)x\theta(x,i)dx\right)=1. \tag{3.16}$$

Thus, (3.3) and (3.4) hold. □

4 Dynamic Programming Equation

We have constructed the stationary distribution $\nu(\cdot,\cdot)$. However, it is generally not easy to approximate the invariant measure. To obtain the optimal ergodic control of dividend payment, we will refer to the dynamic programming equation in (2.14). To solve for (2.14), we will construct a two-component Markov chain to approximate the state process. Then we will rewrite (2.14) by a dynamic programming equation with a Markov chain with transition probabilities.

Before we write the dynamic programming equations, let us recall some results of Markov chains. By using the ergodic theorem for Markov chains in Bertsekas (1987) and Kushner (1972), we can find an auxiliary function $W(x,i,u)$ such that the pair $(W(x,i,u),\gamma(u))$ satisfies

$$W(x,i,u)=\sum_{y}p((x,i),(y,j)|u)W(y,i,u)+u(x,i)x-\gamma(u), \tag{4.1}$$

for each feedback control $u(\cdot)$. $p((x,i),(y,j)|u)$ is the transition probability from a state (x,i) to another state (y,j) under the control $u(\cdot)$.

Define $\bar{\gamma} = \max_u \gamma(u)$, where $u(\cdot) \in \mathcal{U}$. Then there is an auxiliary function $V(x,i)$ such that the pair $(V(x,i),\bar{\gamma})$ satisfies the dynamic programming equation

$$V(x,i) = \max_{u\in\mathcal{U}} \left\{ \sum_y p((x,i),(y,j)|u)V(y,i) + u(x,i)x - \bar{\gamma} \right\}. \tag{4.2}$$

In order to keep $V(x,i)$ from blowing up, (4.2) can be written in a centered form as follows.

$$V(x,i) = \max_{u\in\mathcal{U}} \left\{ \sum_y p((x,i),(y,j)|u)\widetilde{V}(y,i) + u(x,i)x \right\}, \tag{4.3}$$

where

$$\widetilde{V}(y,i) = V(y,i) - V(x_0,i).$$

x_0 is determined such that $\bar{\gamma} = V(x_0,i)$.

Boundary Conditions. For the purpose of the numerical analysis, it is always necessary to consider a compact state space. In our problem, the surplus could potentially grow to arbitrary high level. Our control variable is the dividend payment strategies. When surplus is too high, it is optimal to pay the dividend according to our objective function. Furthermore, the domain of the surplus process is compactified for the computation purpose where a large enough right boundary B was imposed. To be consistent with the reality, it is natural to set a reflecting boundary on the right side. For the left side, the surplus follows a log normal distribution, and is always positive. We will also choose a reflecting boundary on the left side for the computation purpose. Hence, for the boundaries, $V(x,i)$ follows

$$V_x(x,i) = 0. \tag{4.4}$$

5 Numerical Algorithm

Our goal is to design a numerical scheme to approximate $\bar{\gamma}$ in (2.12). In what follows, Section 5.1 will construct an approximating Markov chain in the state space. The discretion of dynamic programming equation is presented in Section 5.2 and the transition probability of the approximating Markov chain is derived.

5.1 Approximating Markov Chain

We will construct a locally consistent Markov chain to approximate the controlled regime-switching diffusion system. The discrete-time controlled Markov chain is so defined that it is locally consistent with (2.9). Note that the state of the process has two components x and α. Hence, in order to use the methodology in Kushner and Dupuis (2001), our approximating Markov chain must have two components: one component delineates the diffusive behavior whereas the other keeps track of the regimes. Let $h > 0$ be a discretization parameter representing the step size. Define $S_h' = \{x : x = kh, k = 0, \pm 1, \pm 2, \dots\}$ and $S_h = S_h' \cap G_h$, where $G_h = (0, B+h)$ and B is an upper bound introduced for numerical computation purpose. Moreover, assume without loss of generality that the boundary point B is an integer multiple of h. Let $\{(\xi_n^h, \alpha_n^h), n < \infty\}$ be a controlled discrete-time Markov chain on $S_h \times \mathcal{M}$ and denote by $p^h((x,i),(y,j)|u^h)$ the transition probability from a state (x,i) to another state (y,j) under the control u^h. We need to define p^h so that the chain's evolution well approximates the local behavior of the controlled regime-switching diffusion process (2.9). We proceed as follows. At any discrete time n, we can either pay a dividend payment as a regular control or a reflection on the boundary. That is, if we put $\Delta\xi^h_n = \xi^h_{n+1} - \xi^h_n$, then

$$\Delta\xi^h_n = \Delta\xi^h_n I_{\{\text{dividend payment at } n\}} + \Delta\xi^h_n I_{\{\text{reflection step on the left at } n\}} + \Delta\xi^h_n I_{\{\text{reflection step on the right at } n\}}. \tag{5.1}$$

The chain and the control will be chosen so that there is exactly one term in (5.1) is nonzero. Denote by $\left\{I^h_n : n = 0, 1, \dots\right\}$ a sequence of control actions, where $I^h_n = 0$, 1, or 2, if we exercise a dividend payment, or reflection on the left or right boundaries at time n, respectively.

If $I^h_n = 0$, then we denote by $u^h_n \subset U$ the random variable that is the dividend payment action for the chain at time n. Let $\tilde{\Delta}t^h(\cdot, \cdot, \cdot) > 0$ be the *interpolation interval* on $S_h \times \mathcal{M} \times U$. Assume $\inf_{x,i,u} \tilde{\Delta}t^h(x, i, u) > 0$ for each $h > 0$ and $\lim_{h\to 0} \sup_{x,i,u} \tilde{\Delta}t^h(x, i, u) \to 0$. If $I^h_n = 1$, or $\xi^h_n = 0$, reflection step on the left boundary is exerted definitely. We require reflection takes the state from 0 to h. That is, if we denote by Δz^h_n the random variable that is the left reflection action size for the chain at time n, then $\Delta\xi^h_n = \Delta z^h_n = h$. If $I^h_n = 2$, or $\xi^h_n = B + h$, reflection step on the right boundary is exerted definitely. We require reflection takes the state from $B + h$ to B. That is, if we denote by Δg^h_n the random variable that is the right reflection action size for the chain at time n, then $\Delta\xi^h_n = -\Delta g^h_n = -h$.

Let $\mathbb{E}^{u,h,0}_{x,i,n}$, $\text{Var}^{u,h,0}_{x,i,n}$ and $\mathbb{P}^{u,h,0}_{x,i,n}$ denote the conditional expectation, variance, and marginal probability given $\{\xi^h_k, \alpha^h_k, u^h_k, I^h_k, k \le n, \xi^h_n = x, \alpha^h_n = i, I^h_n = 0, u^h_n = u\}$, respectively. The sequence $\{(\xi^h_n, \alpha^h_n)\}$ is said to be *locally consistent*, if it satisfies

$$\begin{aligned}
&\mathbb{E}^{u,h,0}_{x,i,n}[\Delta\xi^h_n] = x[\pi(\beta(i) - h(\lambda) - \lambda c(i) + \mu(i)) + \mu(i) - u]\tilde{\Delta}t^h(x, i, u) + o(\tilde{\Delta}t^h(x, i, u)),\\
&\text{Var}^{u,h,0}_{x,i,n}(\Delta\xi^h_n) = (\pi + 1)^2\sigma^2(i)x^2\tilde{\Delta}t^h(x, i, u) + o(\tilde{\Delta}t^h(x, i, u)),\\
&\mathbb{P}^{u,h,0}_{x,i,n}\{\alpha^h_{n+1} = j\} = q_{ij}\tilde{\Delta}t^h(x, i, u) + o(\tilde{\Delta}t^h(x, i, u)), \text{ for } j \ne i,\\
&\mathbb{P}^{u,h,0}_{x,i,n}\{\alpha^h_{n+1} = i\} = 1 + q_{ii}\tilde{\Delta}t^h(x, i, u) + o(\tilde{\Delta}t^h(x, i, u)).\\
&\sup_{n,\omega\in\Omega} |\Delta\xi^h_n| \to 0 \text{ as } h \to 0.
\end{aligned} \tag{5.2}$$

We require the reflections to be "impulsive" or "instantaneous" when $I^h_n = 1$ and $I^h_n = 2$. In other words, the interpolation interval on $S_h \times \mathcal{M} \times U \times \{0, 1, 2\}$ is

$$\Delta t^h(x, i, u, \bar{i}) = \tilde{\Delta}t^h(x, i, u)I_{\{\bar{i}=0\}}, \text{ for any } (x, i, u, \bar{i}) \in S_h \times \mathcal{M} \times U \times \{0, 1, 2\}. \tag{5.3}$$

The sequence u^h is said to be *admissible* if u^h_n is $\sigma\left\{(\xi^h_0, \alpha^h_0), \dots, (\xi^h_n, \alpha^h_n), u^h_0, \dots, u^h_{n-1}\right\}$-adapted and for any $E \in \mathcal{B}(S_h \times \mathcal{M})$, we have

$$\mathbb{P}\left\{(\xi^h_{n+1}, \alpha^h_{n+1}) \in E \middle| \sigma\{(\xi^h_0, \alpha^h_0), \dots, (\xi^h_n, \alpha^h_n), u^h_0, \dots, u^h_n\}\right\} = p^h((\xi^h_n, \alpha^h_n), E|u^h_n),$$

$$\mathbb{P}\left\{(\xi^h_{n+1}, \alpha^h_{n+1}) = (h, i) \middle| (\xi^h_n, \alpha^h_n) = (0, i), \sigma\{(\xi^h_0, \alpha^h_0), \dots, (\xi^h_n, \alpha^h_n), u^h_0, \dots, u^h_n\}\right\} = 1,$$

and

$$\mathbb{P}\left\{(\xi^h_{n+1}, \alpha^h_{n+1}) = (B, i) \middle| (\xi^h_n, \alpha^h_n) = (B + h, i), \sigma\{(\xi^h_0, \alpha^h_0), \dots, (\xi^h_n, \alpha^h_n), u^h_0, \dots, u^h_n\}\right\} = 1.$$

Put

$$t_0^h := 0,\ t_n^h := \sum_{k=0}^{n-1} \Delta t^h(\xi_k^h, \alpha_k^h, u_k^h, I_k^h), \Delta t_k^h = \Delta t^h(\xi_k^h, \alpha_k^h, u_k^h, I_k^h), \quad \text{and } n^h(t) := \max\left\{n : t_n^h \le t\right\}.$$

Then the piecewise constant interpolations, denoted by $(\xi^h(\cdot), \alpha^h(\cdot))$, $u^h(\cdot)$, $z^h(\cdot)$, and $g^h(\cdot)$, are naturally defined as

$$\xi^h(t) = \xi_n^h,\ \alpha^h(t) = \alpha_n^h,\ u^h(t) = u_n^h = u(\xi_n^h),\ z^h(t) = \sum_{k \le n^h(t)} \Delta z_k^h I_{\{I_k^h = 1\}}, g^h(t) = \sum_{k \le n^h(t)} \Delta g_k^h I_{\{I_k^h = 2\}} \tag{5.4}$$

for $t \in [t_n^h, t_{n+1}^h)$. Let $(\xi_0^h, \alpha_0^h) = (x, i) \in S_h \times \mathcal{M}$ and u^h be an admissible control. The cost function for the controlled Markov chain is defined as

$$J_B^h(x, i, u) = \limsup_n \frac{\mathbb{E}_{x,i} \sum_{k=1}^{n-1} u_k^h \xi_k^h \Delta t_k^h}{\mathbb{E}_{x,i} \sum_{k=1}^{n-1} \Delta t_k^h}, \tag{5.5}$$

which is analogous to (2.11) regarding to the definition of interpolation intervals in (5.3). Since $J_B^h(x, i, u)$ does not depend on the initial condition (x, i), we write it as $\gamma^h(u)$. Likewise, we denote

$$\bar{\gamma}^h = \sup_{u^h \text{ admissible}} \gamma^h(u). \tag{5.6}$$

Note that we are considering feedback controls $u(\cdot)$ here. Similarly to ν in (3.7), let $\nu^h(u) = \nu^h(x, u), x \in S_h$ denote the associate invariant measure in the approximating space. Then $\gamma^h(u)$ can be rewritten as

$$\gamma^h(u) = \limsup_n \frac{\mathbb{E}_{x,i} \sum_{k=1}^{n-1} u_k^h \xi_k^h \Delta t_k^h}{\mathbb{E}_{x,i} \sum_{k=1}^{n-1} \Delta t_k^h} = \frac{\sum_{x,i} u(x, i) x \Delta t^h(x, i, u(x, i), 0) \nu^h(x, u)}{\sum_{x,i} \Delta t^h(x, i, u(x, i), 0).\nu^h(x, u)}. \tag{5.7}$$

Since, the time interval of the approximating Markov chain $\Delta t^h(x, i, u(x, i), 0)$ depends on x and u, the invariant measure for the approximating Markov chain needs consider the time spent on each state of the interpolated process. Then, we define a new measure $\omega^h(u) = \omega^h(x, i, u), x \in S_h$ such that

$$\omega^h(x, i, u) = \frac{\Delta t^h(x, i, u(x, i), 0) \nu^h(x, i)}{\sum_{x,i} \Delta t^h(x, i, u(x, i), 0) \nu^h(x, i)}. \tag{5.8}$$

Hence, $\gamma^h(u)$ can be written in a simple form as

$$\gamma^h(u) = \sum_x u(x, i) x \omega^h(x, i, u). \tag{5.9}$$

Let $\mathbb{E}^u_{\omega^h(u)}$ be the expectation for the stationary process under control $u(\cdot)$. In view of (5.7), (5.9) can also be written as

$$\gamma^h(u) = \mathbb{E}^u_{\omega^h(u)} \int_0^1 u(\xi^h(s)) \xi^h(s) ds. \tag{5.10}$$

Remark 5.1. Practically, it is much harder to calculate the invariant measure $\omega^h(u)$ than to calculate the summation $\sum_x u(x, i) x \omega^h(x, i, u)$. By using the iteration method, the convergence speed for computing the value of $\gamma^h(u)$ is much faster than that for computing the invariant measure $\omega^h(u)$. Hence, we focus on the converge of the state process and objective functions instead of the invariant measure itself.

We shall show that $V^h(x,i)$ satisfies the dynamic programming equation:

$$V^h(x,i) = \begin{cases} p^h((x,i),(y,j)|u)V^h(y,j) + (ux-\gamma^h)\Delta t^h(x,i,u,0), & \text{for } x \in S_h, \\ p^h((x,i),(y,j)|u)V^h(y,j), & \text{for } x \in \partial S_h. \end{cases} \tag{5.11}$$

In the actual computing, we use iteration in value space or iteration in policy space together with Gauss-Seidel iteration to solve V^h. The computations will be very involved. In contrast to the usual state space S_h in Kushner and Dupuis (2001), here we need to deal with an enlarged state space $S_h \times \mathcal{M}$ due to the presence of regime switching.

5.2 Discretization

Define the approximation to the first and the second derivatives of $V(\cdot,i)$ by finite difference method in (2.14) using stepsize $h>0$ as:

$$\begin{aligned} &V(x,i) \to V^h(x,i) \\ &V_x(x,i) \to \frac{V^h(x+h,i)-V^h(x,i)}{h} \quad \text{for} \quad x(\pi(\beta(i)-h(\lambda)-\lambda c(i)+\mu(i))+\mu(i)-u)>0, \\ &V_x(x,i) \to \frac{V^h(x,i)-V^h(x-h,i)}{h} \quad \text{for} \quad x(\pi(\beta(i)-h(\lambda)-\lambda c(i)+\mu(i))+\mu(i)-u)<0, \\ &V_{xx}(x,i) \to \frac{V^h(x+h,i)-2V^h(x,i)+V^h(x-h,i)}{h^2}. \end{aligned} \tag{5.12}$$

It leads to, $\forall x \in S_h, i \in \mathcal{M}$,

$$\begin{aligned} \max_{u\in U}\Big\{ &\frac{V^h(x+h,i)-V^h(x,i)}{h}\Big[x(\pi(\beta(i)-h(\lambda)-\lambda c(i)+\mu(i))+\mu(i)-u)\Big]^+ \\ &-\frac{V^h(x,i)-V^h(x-h,i)}{h}\Big[x(\pi(\beta(i)-h(\lambda)-\lambda c(i)+\mu(i))+\mu(i)-u)\Big]^- \\ &+\frac{V^h(x+h,i)-2V^h(x,i)+V^h(x-h,i)}{h^2}\,\frac{(\pi+1)^2\sigma^2(i)x^2}{2}+\sum_j V^h(x,\cdot)q_{ij}+ux-\bar{\gamma}\Big\}=0, \end{aligned} \tag{5.13}$$

where $\Big[x(\pi(\beta(i)-h(\lambda)-\lambda c(i)+\mu(i))+\mu(i)-u)\Big]^+$ and $\Big[x(\pi(\beta(i)-h(\lambda)-\lambda c(i)+\mu(i))+\mu(i)-u)\Big]^-$ are the positive and negative parts of $\Big[x(\pi(\beta(i)-h(\lambda)-\lambda c(i)+\mu(i))+\mu(i)-u)\Big]$, respectively.

For the reflecting boundaries, we choose

$$V_x(x,i) \to \frac{V^h(x,i)-V^h(x-h,i)}{h}. \tag{5.14}$$

Comparing (5.13) and (5.14) with (5.11), we achieve the transition probabilities of $V^h(x,i)$ in the interior of domain as the following:

$$
\begin{aligned}
&p^h((x,i),(x+h,i)|u) = \frac{(\pi+1)^2\sigma^2(i)x^2/2 + h\Big[x(\pi(\beta(i)-h(\lambda)-\lambda c(i)+\mu(i))+\mu(i)-u)\Big]^+}{D},\\
&p^h((x,i),(x-h,i)|u) = \frac{(\pi+1)^2\sigma^2(i)x^2/2 + h\Big[x(\pi(\beta(i)-h(\lambda)-\lambda c(i)+\mu(i))+\mu(i)-u)\Big]^-}{D},\\
&p^h((x,i),(x,j)|u) = \frac{q_{ij}h^2}{D}, \quad \text{for} \quad i \neq j,\\
&p^h(\cdot) = 0, \quad \text{otherwise},\\
&\Delta t^h(x,i,u,2) = \frac{h^2}{D},
\end{aligned}
\tag{5.15}
$$

with

$$
D = (\pi+1)^2\sigma^2(i)x^2 + h|x(\pi(\beta(i)-h(\lambda)-\lambda c(i)+\mu(i))+\mu(i)-u)| - h^2 q_{ii}
$$

being well defined. We also find the transition probability of $V^h(x,i)$ on the boundaries comparing with (5.11) as follows

$$
p^h((x,i),(x+h,i)|u) = 1, \text{ for } x = 0, \tag{5.16}
$$

and

$$
p^h((x,i),(x-h,i)|u) = 1, \text{ for } x = B. \tag{5.17}
$$

6 Convergence of Numerical Approximation

This section focuses on the asymptotic properties of the approximating Markov chain proposed in the last section. The main techniques are methods of weak convergence. To begin with, the technique of time rescaling and the interpolation of the approximation sequences are introduced in Section 6.1. The definition of relax controls is presented in Section 6.2. Section 6.3 deals with weak convergence of $\{\widehat{\xi}^h(\cdot), \widehat{\alpha}^h(\cdot), \widehat{m}^h(\cdot), \widehat{w}^h(\cdot), \widehat{z}^h(\cdot), \widehat{g}^h(\cdot), \widehat{T}^h(\cdot)\}$, a sequence of rescaled process. As a result, a sequence of controlled surplus processes converges to a limit surplus process. Finally Section 6.4 establishes the convergence of the optimal value.

6.1 Interpolation and Rescaling

Based on the approximation Markov chain constructed above, the piecewise constant interpolation is obtained and the appropriate interpolation interval level is chosen. Recalling (5.4), the continuous-time interpolations $(\xi^h(\cdot), \alpha^h(\cdot))$, $u^h(\cdot)$, $g^h(\cdot)$, and $z^h(\cdot)$ are defined. In addition, let $\mathcal{U}^h$ denote the collection of controls, which are determined by a sequence of measurable functions $F_n^h(\cdot)$ such that

$$
u_n^h = F_n^h(\xi_k^h, \alpha_k^h, k \le n; u_k^h, k \le n). \tag{6.1}
$$

Define $\mathcal{D}_t^h$ as the smallest σ-algebra generated by $\{\xi^h(s), \alpha^h(s), u^h(s), g^h(s), z^h(s), s \le t\}$. In addition, $\mathcal{U}^h$ defined by (6.1) is equivalent to the collection of all piecewise constant admissible controls with respect to $\mathcal{D}_t^h$.

For simplicity, let

$$
\begin{aligned}
&b(x,i,u) = x(\pi(\beta(i)-h(\lambda)-\lambda c(i)+\mu(i))+\mu(i)-u),\\
&\sigma(x,i,u) = (\pi+1)\sigma(i)x.
\end{aligned}
$$

Using the representations of regular control, reflection step and the interpolations defined above, (5.1) yields

$$
\begin{aligned}
\xi^h(t) &= x + \sum_{k=0}^{n-1} [\mathbb{E}_k^h \Delta\xi_k^h + (\Delta\xi_k^h - \mathbb{E}_k^h \Delta\xi_k^h)] \\
&= x + \sum_{k=0}^{n-1} b(\xi_k^h, \alpha_k^h, u_k^h)\Delta t^h(\xi_k^h, \alpha_k^h, u_k^h, 0) + \sum_{k=0}^{n-1} (\Delta\xi_k^h - \mathbb{E}_k^h \Delta\xi_k^h) + \varepsilon^h(t) \\
&= x + B^h(t) + M^h(t) + \varepsilon^h(t),
\end{aligned}
\tag{6.2}
$$

where

$$
\begin{aligned}
B^h(t) &= b(\xi_k^h, \alpha_k^h, u_k^h)\Delta t^h(\xi_k^h, \alpha_k^h, u_k^h, 0), \\
M^h(t) &= \sum_{k=0}^{n-1} (\Delta\xi_k^h - \mathbb{E}_k^h \Delta\xi_k^h),
\end{aligned}
$$

and $\varepsilon^h(t)$ is a negligible error satisfying

$$
\lim_{h\to\infty} \sup_{0\le t\le T} E|\varepsilon^h(t)|^2 \to 0 \quad \text{for any} \quad 0 < T < \infty. \tag{6.3}
$$

Also, $M^h(t)$ is a martingale with respect to $\mathcal{D}_t^h$, and its discontinuity goes to zero as $h \to 0$. We attempt to represent $M^h(t)$ similar to the diffusion term in (2.9). Define $w^h(\cdot)$ as

$$
\begin{aligned}
w^h(t) &= \sum_{k=0}^{n-1} (\Delta\xi_k^h - \mathbb{E}_k^h \Delta\xi_k^h)/\sigma(\xi_k^h, \alpha_k^h, u_k^h), \\
&= \int_0^t \sigma^{-1}(\xi^h(s), \alpha^h(s), u^h(s)) dM^h(s).
\end{aligned}
\tag{6.4}
$$

We can now rewrite (6.2) as

$$
\xi^h(t) = x + \int_0^t b(\xi^h(s), \alpha^h(s), u^h(s))ds + \int_0^t \sigma(\xi^h(s), \alpha^h(s), u^h(s))dw^h(s) + \varepsilon^h(t). \tag{6.5}
$$

Now we introduce the rescaling process. The basic idea of rescaling time is to "stretch out" the control and state processes so that they are smoother and the tightness of $g^h(\cdot)$ and $z^h(\cdot)$ can be proved. Define $\Delta\widehat{t}_n^h$ by

$$
\Delta\widehat{t}_n^h = \begin{cases} \Delta t^h & \text{for a diffusion on step } n, \\ |\Delta z_n^h| = h & \text{for a left reflection on step } n, \\ |\Delta g_n^h| = h & \text{for a right reflection on step } n, \end{cases} \tag{6.6}
$$

Define $\widehat{T}^h(\cdot)$ by

$$
\widehat{T}^h(t) = \sum_{i=0}^{n-1} \Delta t^h = t_n^h, \quad \text{for} \ \ t \in [\widehat{t}_n^h, \widehat{t}_{n+1}^h]
$$

Thus, $\widehat{T}^h(\cdot)$ will increase with the slop of unity if an only if a regular control is exerted. In addition, define the rescaled and interpolated process $\widehat{\xi}^h(t) = \xi^h(\widehat{T}^h(t))$, likewise define $\widehat{\alpha}^h(t))$, $\widehat{u}^h(t)$, $\widehat{g}^h(t)$ similarly. The time scale is stretched out by h at the reflection and singular control steps. We can now write

$$
\widehat{\xi}^h(t) = x + \int_0^t b(\widehat{\xi}^h(s), \widehat{\alpha}^h(s), \widehat{u}^h(s))ds + \int_0^t \sigma(\widehat{\xi}^h(s), \widehat{\alpha}^h(s), \widehat{u}^h(s))dw^h(s) + \varepsilon^h(t). \tag{6.7}
$$

6.2 Relaxed Controls

Let $\mathcal{B}(U \times [0,\infty))$ be the σ-algebra of Borel subsets of $U \times [0,\infty)$. An *admissible relaxed control* (or deterministic relaxed control) $m(\cdot)$ is a measure on $\mathcal{B}(U \times [0,\infty))$ such that $m(U \times [0,t]) = t$ for each $t \geq 0$. Given a relaxed control $m(\cdot)$, there is an $m_t(\cdot)$ such that $m(d\phi dt) = m_t(d\phi)dt$. We can define $m_t(B) = \lim_{\delta\to 0} \frac{m(B\times[t-\delta,t])}{\delta}$ for $B \in \mathcal{B}(U)$. With the given probability space, we say that $m(\cdot)$ is an admissible relaxed (stochastic) control for $(w(\cdot), \alpha(\cdot))$ or $(m(\cdot), w(\cdot), \alpha(\cdot))$ is admissible, if $m(\cdot,\omega)$ is a deterministic relaxed control with probability one and if $m(O \times [0,t])$ is $\mathcal{F}_t$-adapted for all $O \in \mathcal{B}(U)$. There is a derivative $m_t(\cdot)$ such that $m_t(\cdot)$ is $\mathcal{F}_t$-adapted for all $O \in \mathcal{B}(U)$.

Given a relaxed control $m(\cdot)$ of $u^h(\cdot)$, we define the derivative $m_t(\cdot)$ such that

$$m^h(K) = \int_{U\times[0,\infty)} I_{\{(u^h,t)\in K\}} m_t(d\phi)dt \tag{6.8}$$

for all $K \in \mathcal{B}(U \times [0,\infty))$, and that for each t, $m_t(\cdot)$ is a measure on $\mathcal{B}(U)$ satisfying $m_t(U) = 1$. For example, we can define $m_t(\cdot)$ in any convenient way for $t = 0$ and as the left-hand derivative for $t > 0$,

$$m_t(O) = \lim_{\delta\to 0} \frac{m(O \times [t-\delta,t])}{\delta},\ \forall O \in \mathcal{B}(U). \tag{6.9}$$

Note that $m(d\phi dt) = m_t(d\phi)dt$. It is natural to define the relaxed control representation $m^h(\cdot)$ of $u^h(\cdot)$ by

$$m_t^h(O) = I_{\{u^h(t)\in O\}},\ \forall O \in \mathcal{B}(U). \tag{6.10}$$

Let $\mathcal{F}_t^h$ be a filtration, which denotes the minimal σ-algebra that measures

$$\{\xi^h(s), \alpha^h(\cdot), m_s^h(\cdot), w^h(s), z^h(s), g^h(s), s \leq t\}. \tag{6.11}$$

Use Γ^h to denote the set of admissible relaxed controls $m^h(\cdot)$ with respect to $(\alpha^h(\cdot), w^h(\cdot))$ such that $m_t^h(\cdot)$ is a fixed probability measure in the interval $[t_n^h, t_{n+1}^h)$ given $\mathcal{F}_t^h$. Then Γ^h is a larger control space containing $\mathcal{U}^h$. Referring to the stretched out time scale, we denote the rescaled relax control as $m_{\widehat{T}^h(t)}(d\psi)$. Define $M_t(O)$ and $M_t^h(d\psi)$ by

$$M_t(O)dt = dw(t)I_{u(t)\in O},\ \forall O \in \mathcal{B}(U)$$
$$M_t^h(d\psi)dt = dw^h(t)I_{u^h(t)\in\mathcal{U}}.$$

Analogously, as an extension of time rescaling, we let

$$\widehat{M}^h_{\widehat{T}^h(t)}(d\psi)d\widehat{T}^h(t) = d\widehat{w}^h(\widehat{T}^h(t))I_{u^h(\widehat{T}^h(t))\in\mathcal{U}}.$$

With the notation of relaxed control given above, we can write (6.5), (6.7), and (5.6) as

$$\xi^h(t) = x + \int_0^t \int_{\mathcal{U}} b(\xi^h(s), \alpha^h(s), \psi)m_s^h(d\psi)ds + \int_0^t \int_{\mathcal{U}} \sigma(\xi^h(s), \alpha^h(s), \psi)M_s^h(d\psi)ds + \varepsilon^h(t), \tag{6.12}$$

$$\begin{aligned}\widehat{\xi}^h(t) = x &+ \int_0^t \int_{\mathcal{U}} b(\widehat{\xi}^h(s), \widehat{\alpha}^h(s), \psi)\widehat{m}^h_{\widehat{T}^h(s)}(d\psi)d\widehat{T}^h(s) \\ &+ \int_0^t \int_{\mathcal{U}} \sigma(\widehat{\xi}^h(s), \widehat{\alpha}^h(s), \psi)\widehat{M}_{\widehat{T}^h(s)}(d\psi)d\widehat{T}^h(s) + \varepsilon^h(t),\end{aligned} \tag{6.13}$$

and

$$\bar{\gamma}^h = \inf_{m^h\in\Gamma^h} \gamma^h(m^h). \tag{6.14}$$

Now we give the definition of existence and uniqueness of weak solution.

Definition 6.1. By a weak solution of (6.12), we mean that there exists a probability space $(\Omega, \mathcal{F}, P)$, a filtration $\mathcal{F}_t$-Wiener process, and process $(x(\cdot), \alpha(\cdot), m(\cdot), w(\cdot))$ such that $w(\cdot)$ is a standard $\mathcal{F}_t$-Wiener process, $\alpha(\cdot)$ is a Markov chain with generator Q and state space $\mathcal{M}$, $m(\cdot)$ is admissible with respect to $x(\cdot)$ is $\mathcal{F}_t$-adapted, and (6.12) is satisfied. For an initial condition (x, i), by the weak sense uniqueness, we mean that the probability law of the admissible process $(\alpha(\cdot), m(\cdot), w(\cdot))$ determines the probability law of solution $(x(\cdot), \alpha(\cdot), m(\cdot), w(\cdot))$ to (6.12), irrespective of probability space.

To proceed, we need some assumptions.

(A1) Let $u(\cdot)$ be an admissible ordinary control with respect to $w(\cdot)$ and $\alpha(\cdot)$, and suppose that $u(\cdot)$ is piecewise constant and takes only a finite number of values. For each initial condition, there exists a solution to (6.12) where $m(\cdot)$ is the relaxed control representation of $u(\cdot)$. This solution is unique in the weak sense.

6.3 Convergence of a Sequence of Surplus Processes

In this section, we will deal with the convergence of the approximation sequence to the regime-switching process and the surplus process. We will derive one lemma and three theorems, whose proof are provided in the Appendix.

Lemma 6.2. *Using the transition probabilities $\{p^h(\cdot)\}$ defined in* (5.15)*, the interpolated process of the constructed Markov chain $\{\widehat{\alpha}^h(\cdot)\}$ converges weakly to $\widehat{\alpha}(\cdot)$, the Markov chain with generator $Q = (q_{\ell\iota})$.*

Theorem 6.3. *Let the approximating chain $\{\xi_n^h, \alpha_n^h, n < \infty\}$ constructed with transition probabilities defined in* (5.15) *be locally consistent with* (2.9)*, $m^h(\cdot)$ be the relaxed control representation of $\{u_n^h, n < \infty\}$, $(\xi^h(\cdot), \alpha^h(\cdot))$ be the continuous-time interpolation defined in* (5.4)*, and $\{\widehat{\xi}^h(\cdot), \widehat{\alpha}^h(\cdot), \widehat{m}^h(\cdot), \widehat{w}^h(\cdot), \widehat{z}^h(\cdot), \widehat{g}^h(\cdot), \widehat{T}^h(\cdot)\}$ be the corresponding rescaled processes. Then $\{\widehat{\xi}^h(\cdot), \widehat{\alpha}^h(\cdot), \widehat{m}^h(\cdot), \widehat{w}^h(\cdot), \widehat{z}^h(\cdot), \widehat{g}^h(\cdot), \widehat{T}^h(\cdot)\}$ is tight.*

Theorem 6.4. *Let $\{\widehat{x}(\cdot), \widehat{\alpha}(\cdot), \widehat{m}(\cdot), \widehat{w}(\cdot), \widehat{z}(\cdot), \widehat{g}(\cdot), \widehat{T}(\cdot)\}$ be the limit of weakly convergent subsequence of $\{\widehat{\xi}^h(\cdot), \widehat{\alpha}^h(\cdot), \widehat{m}^h(\cdot), \widehat{w}^h(\cdot), \widehat{z}^h(\cdot), \widehat{g}^h(\cdot), \widehat{T}^h(\cdot)\}$. $w(\cdot)$ is a standard $\mathcal{F}_t$-Wiener process, and $m(\cdot)$ is admissible. Let $\widehat{\mathcal{F}}_t$ be the σ-algebra generated by $\{\widehat{\xi}^h(\cdot), \widehat{\alpha}^h(\cdot), \widehat{m}^h(\cdot), \widehat{w}^h(\cdot), \widehat{z}^h(\cdot), \widehat{g}^h(\cdot), \widehat{T}^h(\cdot)\}$. Then $\widehat{w}(t) = w(\widehat{T}(t))$ is an $\widehat{\mathcal{F}}_t$-martingale with quadratic variation $\widehat{T}(t)$. The limit processes satisfy*

$$\widehat{x}(t) = x + \int_0^t \int_{\mathcal{U}} b(\widehat{x}(s), \widehat{\alpha}(s), \psi) \widehat{m}^h_{\widehat{T}(s)}(d\psi) d\widehat{T}(s) + \int_0^t \int_{\mathcal{U}} \sigma(\widehat{x}(s), \widehat{\alpha}(s), \psi) \widehat{M}_{\widehat{T}(s)}(d\psi) d\widehat{T}(s). \quad (6.15)$$

Theorem 6.5. *For $t < \infty$, define the inverse*

$$\mathcal{T}(t) = \inf\{s : \widehat{T}(s) > t\}.$$

Then $\mathcal{T}(t)$ is right continuous and $\mathcal{T}(t) \to \infty$ as $t \to \infty$ w.p. 1. For any process $\widehat{\varphi}(\cdot)$, define the rescaled process $\varphi(\cdot)$ by $\varphi(t) = \widehat{\varphi}(T(t))$. Then, $w(\cdot)$ is a standard $\mathcal{F}_t$-Wiener process and (2.9) *holds.*

6.4 Convergence of the Optimal Value

To prove the convergence of the optimal value of the objective function, we proceed to find a comparison ε-optimal control.

Lemma 6.6. *For each $\varepsilon > 0$, there exists a continuous feedback control $u^\varepsilon(\cdot)$ that is ε-optimal to all admissible controls. The solution to* (2.9) *is unique in weak sense and has a unique invariant measure under this ε-optimal control.*

Proof. The existence of a smooth ε-optimal can be guaranteed by modifying the method in Kushner (1978) for our formulation. □

Theorem 6.7. *Assume the conditions of Theorem* 6.4 *and Theorem* 6.5 *are satisfied. Then as $h \to 0$,*

$$\bar{\gamma}^h(x,i) \to \bar{\gamma}. \tag{6.16}$$

Proof. First, to prove

$$\bar{\gamma}^h(x,i) \le \bar{\gamma}. \tag{6.17}$$

Let $\tilde{u}(\cdot)$ be the optimal control and $\tilde{m}^h(\cdot)$ be the relaxed control representation of $\tilde{u}^h(\cdot)$. Then, $\bar{\gamma}^h(x,i) = \gamma^h(\tilde{u}^h)$. Hence, in view of (5.10),

$$\begin{aligned}
\bar{\gamma}^h(x,i) &= \mathbb{E}^{\tilde{u}^h} \int_0^1 u(\xi^h(s))\xi^h(s)ds \\
&= \mathbb{E}^{\tilde{u}^h} \int_0^1 \int_{\mathcal{U}} \xi^h(s)\psi \tilde{m}_s^h(d\psi)ds \\
&\to \mathbb{E}^{\tilde{m}} \int_0^1 \int_{\mathcal{U}} x(s)\psi \tilde{m}_s(d\psi)ds \\
&= \lim_T \frac{1}{T}\mathbb{E}^{\tilde{m}} \int_0^T \int_{\mathcal{U}} \psi \tilde{m}_s(d\psi)ds \\
&= \gamma(\tilde{m}) \\
&\le \bar{\gamma},
\end{aligned} \tag{6.18}$$

where $\gamma(\tilde{m})$ is the optimal value of the performance function for the limit stationary process.

On the other hand, from Lemma 6.6, we have ε-optimal control u^ε such that

$$\begin{aligned}
\bar{\gamma}^h(x,i) &\ge \gamma^h(u^\varepsilon, i) \\
&= \mathbb{E}^{u^\varepsilon} \int_0^1 u^\varepsilon(\xi^h(s))\xi^h(s)ds \\
&\to \mathbb{E}^{u^\varepsilon} \int_0^1 u^\varepsilon(x(s))x(s)ds \\
&= \lim_T \frac{1}{T}\mathbb{E}^{u^\varepsilon} \int_0^T u^\varepsilon(x(s))x(s)ds \\
&= \gamma(u^\varepsilon, i) \\
&\ge \bar{\gamma} - \varepsilon.
\end{aligned} \tag{6.19}$$

Combining (6.18) and (6.19) yields (6.16). □

7 Numerical Examples and Further Remarks

7.1 Numerical Examples

This section is devoted to several examples. For simplicity, we consider the case that the discrete event has two states. That is, the continuous-time Markov chain has two states. By using value iteration methods, we numerically solve the optimal control problems.

Based on the algorithm constructed above, we carry out the computation by value iterations. For $n \in Z^+$ and $i \in \mathcal{M}$, define the vectors

$$V_n^h = \{V_n^h(h,1), V_n^h(2h,1), \dots, V_n^h(B,1), \dots V_n^h(h,n_0), V_n^h(2h,m), \dots, V_n^h(B,m)\}$$
$$V^h = \{V_n(h,1), V_n(2h,1), \dots, V_n(B,1), \dots V_n(h,m), V_n(2h,m), \dots, V_n(B,m)\}.$$

Using the method of value iteration, we obtain $V_n(x_0, i)$. The numerical experiments demonstrate that $V_n(x_0, i) \to \bar{\gamma}$ as $n \to \infty$. The procedure is as follows.

1. Set $n = 0$. $\forall x \in S_h$ and $i \in \mathcal{M}$, we set the initial value $V_0^h(x,i) = \widetilde{V}_0^h(x) = 0$.
2. Choose a x_0. Find improved values $V_{n+1}^h(x,i)$ by iteration and record the corresponding optimal control.

$$V_{n+1}^h(x,i) = \max_{u \in \mathcal{U}} \Big[\sum_{(y,j)} (p^h((x,i),(y,j))|u) \widetilde{V}_n^h(y,i) + ux \Big],$$

$$\widetilde{V}_n^h(x,i) = V_n^h(x,i) - V_n^h(x_0,i)$$

3. If $|V_{n+1}^h - V_n^h| >$ tolerance, then $n \to n+1$ and go to step 2; else the iteration stops.

The continuous-time Markov chain $\alpha(t)$ representing the discrete event state has the generator

$$Q = \begin{pmatrix} -10 & 10 \\ 800 & -800 \end{pmatrix},$$

and takes values in $\mathcal{M} = \{1, 2\}$. The claim severity distribution follows exponential distribution with density function $f(y) = ae^{-ay}$ where $a = 0.1$. The premium rate depends on the discrete state with $\beta(1) = 0.02$ and $\beta(2) = 0.06$. The dividend rate $u(t)$ taking values in $[0, 1]$ is the control. Corresponding to the different discrete states, the yield rate of the asset is $\mu(1) = 0.1$ and $\mu(2) = 0.02$. The volatility of the financial market $\sigma(\alpha(t))$ is valued as $\sigma(1) = 0.05$ and $\sigma(2) = 0.1$. The claim rates in different regimes are set as $c(1) = 0.01$ and $c(2) = 0.1$. Hence, we are considering two insurance market modes to represent the insurance cycle. Market mode 1 represents a "soft" market, where the investment return is high and the premium rate is low. While market mode 2 represents a "hard" market, where the investment return is low and the premium rate is high. Obviously, it is much easier for insurance companies when the market is in mode 1. The insurance company is more likely to expand its business and write more policies. Then the liability ratio is higher. Hence, we set $\pi(1)$=0.8. For market mode 2, the insurance and financial market is much harder. The insurance company will preserve sufficient surplus and write less policies to pay for the future claims. Then the liability ratio is lower. Then, we set $\pi(1)$=0.2. To compute the optimal average dividend payment, we choose the value iteration and impose the upper bound of the computation interval of surplus as $B = 50$.

Furthermore, note that the x_0 is arbitrarily chosen to initiate the algorithm. Theoretically, different x_0 is supposed to lead to the same $\bar{\gamma}$. In practical computation, there is inevitable computational errors in calculating the optimal value of γ. The optimal values are the average of convergent values of $V_n(x_0)$ in all available x_0's. To show the stability of the convergence with different x_0's, we plot the values $V_n(x_0)$ with respect to x_0 when the iteration stops in Figure 7.1.

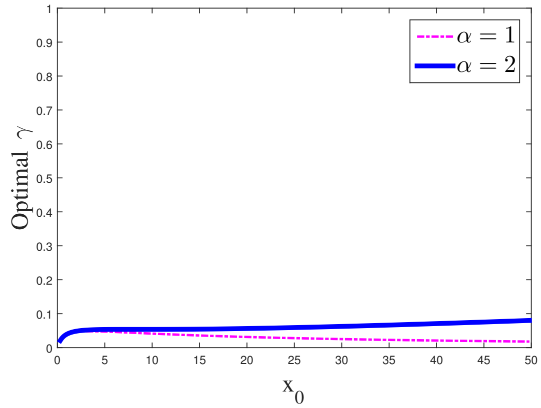

Figure 7.1: Optimal γ versus initial status

From Figure 7.1, it is shown that $V_n(x_0)$ is fluctuating within the range $[0, 0.1]$. After a small hike when x_0 is small, the convergence value of $V_n(x_0)$ is flat and stable when x_0 is bigger. According to the stationary of the process, the average of the convergence value of $V_n(x_0)$ is a good approximation to the optimal long-term average dividend payment. The average of the values of blue dots is 0.06, and average of the values of red dots is 0.03. The optimal ergodic control of dividend payment can be approximated by the mean as 0.045.

7.2 Further Remarks

This work focused on finding the optimal ergodic dividend payment strategies of an insurance company with a long-term goal, taking into account the reinsurance policies. The parameters in the model including premium rate, return rate of the assets and claim rate, depend on the state of economy, which is described by a finite state continuous-time Markov chain. Incorporating the impact of reinsurance on the financial status of the insurance companies, we aimed to maximize the long-run average dividend payment in an infinite time horizon. A generalized stationary diffusion process of surplus is presented. The invariant measure is constructed and the optimal value is obtained correspondingly. By using the dynamic programming approach, we derive the associated system of HJB equations. However, due to the regime-switching, approximating the invariant measure is very difficult. Then we design a numerical scheme to approximate the optimal ergodic dividend payment strategy directly. A two-component discrete-time controlled Markov chain is constructed to approximate the controlled regime-switching diffusion process yielding approximation to the optimal value. Convergence of the approximation algorithms is provided. The economic insights shown in the example provide guidance for decision makers in government or industries to manage the leverage level and dividend policies.

In future studies, the techniques of constructing invariant measures and approximating Markov chain can be to extended to a variety of optimization problems of risk-sensitive controls for ergodic processes, where the objective is to maximize/minimize various performance functions over long term. Although the specific aim in this paper was devoted to developing the optimal long-term insurance policies, the methods can be readily adopted to treat other optimal control problems with a long-run average aim and regime-switching diffusion formulation. Future effort may also be devoted to other variant of related game problems with long-term goal objective functions.

Acknowledgments

The research of Zhuo Jin was supported in part by Faculty Research Grant by The University of Melbourne. The research of H. Yang was supported in part by Research Grants Council of the Hong Kong Special Administrative Region (project No. HKU 705313P) and Society of Actuaries' Centers of Actuarial Excellence Research Grant. The research of G. Yin was supported in part by the National Science Foundation under DMS-1710827.

A Appendix

A.1 Proof of Lemma 6.2

Proof. It can be seen that $\alpha^h(\cdot)$ is tight. The proof can be obtained similar to Theorem 3.1 in Yin et al. (2003). Then so is $\widehat{\alpha}^h(\cdot)$ due to the rescaled time. □

A.2 Proof of Theorem 6.3

Proof. In view of Lemma 6.2, $\{\widehat{\alpha}^h(\cdot)\}$ is tight. The sequence $\{\widehat{m}^h(\cdot)\}$ is tight since its range space is compact. Let $T < \infty$, and let τ_h be an $\mathcal{F}_t$-stopping time which is not larger than T. Then for $\delta > 0$,

$$\mathbb{E}^{u^h}_{\tau_h}(w^h(\tau_h+\delta) - w^h(\tau_h))^2 = \delta + \varepsilon_h, \tag{A.1}$$

where $\varepsilon_h \to 0$ uniformly in τ_h. Taking $\limsup_{h\to 0}$ followed by $\lim_{\delta\to 0}$ yield the tightness of $\{w^h(\cdot)\}$. Similar to the argument of $\alpha^h(\cdot)$, the tightness of $\widehat{w}^h(\cdot)$ is obtained. Furthermore, following the definition of "stretched out" timescale,

$$\begin{aligned} |\widehat{z}^h(\tau_h+\delta) - \widehat{z}^h(\tau_h)| &\le |\delta| + O(h), \\ |\widehat{g}^h(\tau_h+\delta) - \widehat{g}^h(\tau_h)| &\le |\delta| + O(h). \end{aligned}$$

Thus $\{\widehat{z}^h(\cdot), \widehat{g}^h(\cdot)\}$ is tight. These results and the boundedness of $b(\cdot)$ implies the tightness of $\{\xi^h(\cdot)\}$. Thus, $\{\widehat{\xi}^h(\cdot), \widehat{\alpha}^h(\cdot), \widehat{u}^h(\cdot), \widehat{w}^h(\cdot), \widehat{z}^h(\cdot), \widehat{g}^h(\cdot), \widehat{T}^h(\cdot)\}$ is tight. □

Since $\{\widehat{x}^h(\cdot), \widehat{\alpha}^h(\cdot), \widehat{m}^h(\cdot), \widehat{w}^h(\cdot), \widehat{z}^h(\cdot), \widehat{g}^h(\cdot), \widehat{T}^h(\cdot)\}$ is tight, we can extract a weakly convergent subsequence denoted by $\{\widehat{\xi}(\cdot), \widehat{\alpha}(\cdot), \widehat{m}(\cdot), \widehat{w}(\cdot), \widehat{z}(\cdot), \widehat{g}(\cdot), \widehat{T}(\cdot)\}$. Also, the paths of $\{\widehat{x}(\cdot), \widehat{\alpha}(\cdot), \widehat{m}(\cdot), \widehat{w}(\cdot), \widehat{z}(\cdot), \widehat{g}(\cdot), \widehat{T}(\cdot)\}$ are continuous w.p. 1. □

A.3 Proof of Theorem 6.4

Proof. For $\delta > 0$, define the process $l(\cdot)$ by $l^{h,\delta}(t) = l^h(n\delta), t \in [n\delta, (n+1)\delta)$. Then, by the tightness of $\{\widehat{\xi}^h(\cdot), \widehat{\alpha}^h(\cdot)\}$, (6.13) can be rewritten as

$$\begin{aligned} \widehat{\xi}^h(t) = x &+ \int_0^t \int_{\mathcal{U}} b(\widehat{\xi}^h(s), \widehat{\alpha}^h(s), \psi) \widehat{m}^h_{\widehat{T}^h(s)}(d\psi) d\widehat{T}^h(s) \\ &+ \int_0^t \int_{\mathcal{U}} \sigma(\widehat{\xi}^{h,\delta}(s), \widehat{\alpha}^{h,\delta}(s), \psi) \widehat{M}_{\widehat{T}^h(s)}(d\psi) d\widehat{T}^h(s) + \varepsilon^{h,\delta}(t), \end{aligned} \tag{A.2}$$

where

$$\lim_{\delta\to 0} \limsup_{h\to 0} \mathbb{E}|\varepsilon^{h,\delta}(t)| = 0. \tag{A.3}$$

If we can verify $\widehat{w}(\cdot)$ is an $\widehat{\mathcal{F}}_t$-martingale, then (6.15) could be obtained by taking limits in (A.2). To characterize $w(\cdot)$, let $t > 0$, $\delta > 0$, p, q, $\{t_k : k \le p\}$ be given such that $t_k \le t \le t+s$ for all $k \le p$, $\psi_j(\cdot)$, for $j \le q$, is real-valued and continuous functions on $U \times [0, \infty)$ having compact support for all $j \le q$. Define

$$(\psi_j, \widehat{m})_t = \int_0^t \int_{\mathcal{U}} \psi_j(\psi, s) \widehat{m}^h_{\widehat{T}(s)}(d\psi) d\widehat{T}(s). \tag{A.4}$$

Let $S(\cdot)$ be a real-valued and continuous function of its arguments with compact support. By (6.4), $w^h(\cdot)$ is an $\mathcal{F}_t$-martingale. In view of the definition of $\widehat{w}(t)$, we have

$$\mathbb{E}S(\widehat{\xi}^h(t_k), \widehat{\alpha}^h(t_k), \widehat{w}^h(t_k), (\psi_j, m^h)_{t_k}, \widehat{z}^h(t_k), \widehat{g}^h(t_k), j \le q, k \le p)[\widehat{w}^h(t+s) - \widehat{w}^h(t)] = 0. \tag{A.5}$$

By using the Skorohod representation and the dominant convergence theorem, letting $h \to 0$, we obtain

$$\mathbb{E}S(\widehat{\xi}^h(t_k), \widehat{\alpha}^h(t_k), \widehat{w}^h(t_k), (\psi_j, m^h)_{t_k}, \widehat{z}^h(t_k), \widehat{g}^h(t_k), j \le q, k \le p)[\widehat{w}(t+s) - \widehat{w}(t)] = 0. \tag{A.6}$$

Since $\widehat{w}(\cdot)$ has continuous sample paths, (A.6) implies that $\widehat{w}(\cdot)$ is a continuous $\mathcal{F}_t$-martingale. On the other hand, since

$$\mathbb{E}[((\widehat{w}^h(t+\delta))^2 - (\widehat{w}^h(t))^2] = \mathbb{E}[(\widehat{w}^h(t+\delta) - \widehat{w}^h(t))^2] = \widehat{T}(t+s) - \widehat{T}(t), \tag{A.7}$$

by using the Skorohod representation and the dominant convergence theorem together with (A.7), we have

$$\begin{aligned} &\mathbb{E}S(\widehat{\xi}^h(t_k), \widehat{\alpha}^h(t_k), \widehat{w}^h(t_k), (\psi_j, m^h)_{t_k}, \widehat{z}^h(t_k), \widehat{g}^h(t_k), j \le q, k \le p) \\ &\qquad [\widehat{w}^2(t+\delta) - \widehat{w}^2(t) - (\widehat{T}(t+s) - \widehat{T}(t))] = 0. \end{aligned} \tag{A.8}$$

The quadratic variation of the martingale $\widehat{w}(t)$ is $\Delta\widehat{T}$. Consequently, $\widehat{w}(\cdot)$ is an $\widehat{\mathcal{F}}_t$-Wiener process.

Let $h \to 0$, by using the Skorohod representation, we obtain

$$\mathbb{E}\left|\int_0^t \int_{\mathcal{U}} b(\widehat{\xi}^h(s), \widehat{\alpha}^h(s), \psi) \widehat{m}^h_{\widehat{T}^h(s)}(d\psi) d\widehat{T}^h(s) - \int_0^t \int_{\mathcal{U}} b(\widehat{x}(s), \widehat{\alpha}(s), \psi) \widehat{m}^h_{\widehat{T}(s)}(d\psi) d\widehat{T}(s)\right| \to 0 \tag{A.9}$$

uniformly in t with probability one. On the other hand, $\{\widehat{m}^h(\cdot)\}$ converges in the compact weak topology, that is, for any bounded and continuous function $\psi(\cdot)$ with compact support,

$$\int_0^\infty \int_{\mathcal{U}} \psi(\psi, s) \widehat{m}^h_{\widehat{T}^h(s)}(d\psi) d\widehat{T}^h(s) \to \int_0^\infty \int_{\mathcal{U}} \psi(\psi, s) \widehat{m}_{\widehat{T}(s)}(d\psi) d\widehat{T}(s). \tag{A.10}$$

Again, the Skorohod representation implies that as $h \to 0$,

$$\int_0^t \int_{\mathcal{U}} b(\widehat{\xi}^h(s), \widehat{\alpha}^h(s), \psi) \widehat{m}^h_{\widehat{T}^h(s)}(d\psi) d\widehat{T}^h(s) \to \int_0^t \int_{\mathcal{U}} b(\widehat{x}(s), \widehat{\alpha}(s), \psi) \widehat{m}_{\widehat{T}(s)}(d\psi) d\widehat{T}(s) \tag{A.11}$$

uniformly in t with probability one on any bounded interval.

In view of (A.2), since $\xi^{h,\delta}(\cdot)$ and $\alpha^{h,\delta}(\cdot)$ are piecewise constant functions,

$$\int_0^t \int_{\mathcal{U}} \sigma(\widehat{\xi}^{h,\delta}(s), \widehat{\alpha}^{h,\delta}(s), \psi) \widehat{M}_{\widehat{T}^h(s)}(d\psi) d\widehat{T}^h(s) \to \int_0^t \int_{\mathcal{U}} \sigma(\widehat{x}^\delta(s), \widehat{\alpha}^\delta(s), \psi) \widehat{M}_{\widehat{T}(s)}(d\psi) d\widehat{T}(s) \tag{A.12}$$

as $h \to 0$. Combining (A.4)-(A.12), we have

$$\begin{aligned}\widehat{x}(t) = x &+ \int_0^t \int_{\mathcal{U}} b(\widehat{x}(s), \widehat{\alpha}(s), \psi)\widehat{m}^h_{\widehat{T}(s)}(d\psi)d\widehat{T}(s) \\ &+ \int_0^t \int_{\mathcal{U}} \sigma(\widehat{x}^\delta(s), \widehat{\alpha}^\delta(s), \psi)\widehat{M}_{\widehat{T}(s)}(d\psi)d\widehat{T}(s) + \varepsilon^\delta(t),\end{aligned} \tag{A.13}$$

where $\lim_{\delta\to 0} \mathbb{E}|\varepsilon^\delta(t)| = 0$. Finally, taking limits in the above equation as $\delta \to 0$, (6.15) is obtained. □

A.4 Proof of Theorem 6.5

Proof. Since $\widehat{T}(t) \to \infty$ w.p. 1 as $t \to \infty$, $\mathcal{T}(t)$ exists for all t and $\mathcal{T}(t) \to \infty$ as $t \to \infty$ w.p. 1. Similar to (A.6) and (A.8),

$$\mathbb{E}S(\xi^h(t_k), \alpha^h(t_k), w^h(t_k), (\psi_j, m^h)_{t_k}, z^h(t_k), g^h(t_k), j \le q, k \le p) \times [w(t+s) - w(t)] = 0.$$

$$\begin{aligned}&\mathbb{E}S(\xi^h(t_k), \alpha^h(t_k), w^h(t_k), (\psi_j, m^h)_{t_k}, z^h(t_k), g^h(t_k), j \le q, k \le p) \\ &\qquad\qquad \times[w^2(t+\delta) - w^2(t) - (\mathcal{T}(t+s) - \mathcal{T}(t))] = 0.\end{aligned}$$

Thus, we can verify $w(\cdot)$ is an $\mathcal{F}_t$-Wiener process. A rescaling of (6.15) yields

$$x(t) = x + \int_0^t \int_{\mathcal{U}} b(x(s), \alpha(s), \psi)m_s(d\psi)ds + \int_0^t \int_{\mathcal{U}} \sigma(x(s), \alpha(s), \psi)M_s(d\psi)ds. \tag{A.14}$$

In other words, (2.9) holds. □

References

Alvarez, L. H. R. and Lempa, J. (2008). On the optimal stochastic impulse control of linear diffusions. *SIAM Journal on Control and Optimization*, 47(2):703–732.

Avram, F., Palmowski, Z., and Pistorius, M. R. (2007). On the optimal dividend problem for a spectrally negative lévy process. *The Annals of Applied Probability*, 17(1):156–180.

Azcue, P. and Muler, N. (2010). Optimal investment policy and dividend payment strategy in an insurance company. *The Annals of Applied Probability*, 20(4): 1253–1302.

Bai, L. and Paulsen, J. (2012). On non-trivial barrier solutions of the dividend problem for a diffusion under constant and proportional transaction costs. *Stochastic Processes and their Applications*, 122(12):4005 – 4027.

Bertsekas, D.P. (1987). *Dynamic Programming: Deterministic and Stochastic Models*. Prentice-Hall, Englewood Cliffs, NJ

Bielecki, T. R., and S. R. Pliska (1999). Risk Sensitive Dynamic Asset Management. *Appl. Math. Optim.*, 39:337–360.

Bielecki, T. R., and S. R. Pliska (2003). "Economic Properties of the Risk Sensitive Criterion for Portfolio Management". *Review of Accounting and Finance*, Vol. 2, 2:3–17.

De Finetti, B. (1957). Su unimpostazione alternativa della teoria collettiva del rischio, *Transactions of the XVth International Congress of Actuaries*, 2 (1957): 433–443.

Fleming, W. H., and McEneaney, W. M. (1995). Risk-Sensitive Control on an Infinite Time Horizon. *SIAM J. Control Optim.*, 33:1881–1915.

Fleming, W. H., and Sheu, S. J. (2000). Risk-Sensitive Control and an Optimal Investment Model. *Mathematical Finance*, Vol. 10, 2:197–213

Fleming, W. and Soner, H. (2006). *Controlled Markov Processes and Viscosity Solutions*, volume 25 of *Stochastic Modelling and Applied Probability.* Springer-Verlag, New York, NY, second edition.

Guo, X., Liu, J., and Zhou, X. Y. (2004). A constrained non-linear regular-singular stochastic control problem, with applications. *Stochastic Processes and their Applications*, 109(2):167 – 187.

He, L. and Liang, Z. (2008). Optimal financing and dividend control of the insurance company with proportional reinsurance policy. *Insurance: Mathematics and Economics*, 42(3):976 – 983.

Jin, Z., Yang, H., Yin, G. (2013). Numerical methods for optimal dividend payment and investment strategies of regime-switching jump diffusion models with capital injections, *Automatica,* 49(8): 2317–2329.

Jin, Z., Yang, H., Yin, G. (2015). Optimal debt ratio and dividend payment strategies with reinsurance, *Insurance Mathematics and Economics*, 64:351-363.

Jin, Z., Yin, G., and Yang, H. (2011). Numerical methods for dividend policy of regime-switching jump-diffusion models. *Mathematical Control and Related Fields*, 1(1):21–40.

Kushner, H. J. (1972). *Introduction to Stochastic Control Theory.* Holt, Rinehart and Winston, New York

Kushner, H. J. (1978). Optimality conditions for the average cost per unit time problem with a diffusion model, *SIAM J. Control Optimization,* 16:330-346

Kushner, H. and Dupuis, P. (2001). *Numerical Methods for Stochstic Control Problems in Continuous Time*, volume 24 of *Stochastic Modelling and Applied Probability.* Springer, New York, second edition.

Loeffen, R. L. and Renaud, J.-F. (2010). De Finetti's optimal dividends problem with an affine penalty function at ruin. *Insurance: Mathematics and Economics*, 46(1):98 – 108. Gerber-Shiu Functions / Longevity risk and capital markets.

Løkka, A. and Zervos, M. (2008). Optimal dividend and issuance of equity policies in the presence of proportional costs. *Insurance: Mathematics and Economics*, 42:954–961.

Meng, H. and Siu, T. K. (2011). Impulse control of proportional reinsurance with constraints. *International Journal of Stochastic Analysis*, 2011:Article ID 190603.

Pham, Huyên (2003). A large deviations approach to optimal long term investment. *Finance Stochast.* 7:169–195.

Sotomayor, L. and Cadenillas, A. (2011) Classical, singular, and impulse stochastic control for the optimal dividend policy when there is regime switching. *Insurance: Mathematics and Economics*, 48(3), 344–354

Wei, F., Wu, L., and Zhou, D. (2014). Optimal control problem for an insurance surplus model with debt liability. *Mathematical Methods in the Applied Sciences*, 37(11):1652–1667.

Wei, J., Yang, H. and Wang, R. (2010). Classical and impulse control for the optimization of dividend and proportional reinsurance policies with regime switching. *Journal of Optimization Theory and Applications*, 147(2).

Yin, G., Zhang, Q., and Badowski, G. (2003). Discrete-time singularly perturbed Markov chains: Aggregation, occupation measures, and switching diffusion limit. *Advances in Applied Probability*, 35: 449–476.

Yin, G. and Zhu, C. (2010). *Hybrid Switching Diffusions: Properties and Applications*. Springer, New York

Zhu, J. (2014). Dividend optimization for a regime-switching diffusion model with restricted dividend rates. *ASTIN Bulletin*, 44:459–494.

Zhu, J. and Chen, F. (2013). Dividend optimization for regime-switching general diffusions. *Insurance: Mathematics and Economics*, 53(2):439 – 456.

Zhu, J. and Yang, H. (2015). Optimal financing and dividend distribution in a general diffusion model with regime switching. *Journal of Applied Probablility*, In press.