



Brief paper

Mean–variance portfolio selection with dynamic attention behavior in a hidden Markov model[☆]Yu Zhang^a, Zhuo Jin^b, Jiaqin Wei^{a,*}, George Yin^c^a Key Laboratory of Advanced Theory and Application in Statistics and Data Science-MOE, School of Statistics, East China Normal University, Shanghai 200062, China^b Department of Actuarial Studies and Business Analytics, Macquarie University, 2109, NSW, Australia^c Department of Mathematics, University of Connecticut, Storrs, CT 06269-1009, USA

ARTICLE INFO

Article history:

Received 28 June 2021

Received in revised form 23 May 2022

Accepted 16 August 2022

Available online xxxx

Keywords:

Mean–variance portfolio selection

Dynamic attention behavior

Extended HJB equation

Markov chain approximation

Hidden Markov model

ABSTRACT

In this paper, we study closed-loop equilibrium strategies for mean–variance portfolio selection problems in a hidden Markov model with dynamic attention behavior. In addition to the investment strategy, the investor's attention to news is introduced as a control of the accuracy of the news signal process. The objective is to find equilibrium strategies by numerically solving an extended HJB equation by using Markov chain approximation method. An iterative algorithm is constructed and its convergence is established. Numerical examples are provided to illustrate the results.

© 2022 Elsevier Ltd. All rights reserved.

1. Introduction

Since the introduction of mean–variance (MV, for short) portfolio selection problem by Markowitz (1952), much progress has been made. Various extensions have been investigated including the dynamic MV problem in multi-period model and continuous-time models; see, e.g., Li and Ng (2000) and Zhou and Li (2000). Certain limitations of a single diffusion and the existence of 'regimes' that switch among themselves in the underlying market are realized, a system commonly referred to as the regime-switching (RS, for short) model was used to discuss MV problems (see, e.g., Zhou and Yin (2003)). Usually, some states of the market cannot be observed by an investor. In the financial market with partially observable states, hidden Markov chain is often used to describe the evolution of the unobservable market states. The portfolio optimization problem in the hidden Markov RS financial market has been investigated by Elliott, Siu, and Badescu (2010) and Yang, Yin, and Zhang (2015).

It is well-known that the dynamic MV optimization problem is time-inconsistent in the sense that Bellman's optimality principle does not hold. In Strotz (1955), where a time-inconsistent

problem within a game theoretic framework was studied, the author viewed the time-inconsistent optimization problem as a multi-person game and looked for a subgame perfect Nash equilibrium point. The precise definition of the equilibrium concept in continuous time within the class of closed-loop strategies was provided for the first time in Ekeland and Lazrak (2006) and Ekeland and Pirvu (2008). They investigated the optimal investment and consumption problem with hyperbolic discounting in continuous-time deterministic and stochastic models, respectively. Recently, there has been increasing interests in finding the so-called equilibrium strategies for the time-inconsistent optimization problems; see Björk and Murgoci (2010), Hu, Jin, and Zhou (2012), Wang, Jin, and Wei (2019), Wei and Wang (2017), among others.

In most of the existing literature on hidden Markov model, it is assumed that the accuracy of signal collection is fixed in the financial market, and investors acquire information passively in the sense that they do not control the quality of the information they collect. However, it is reasonable to introduce costly information-acquisition problems, which can help investors obtain accurate information, but at the expense of decreasing her current wealth. In Detemple and Kihlstrom (1987), the author examined the demand for information and derived the equilibrium price of information. An investor who faces uncertainty about stock-return predictability was considered in Xia (2001). The authors solved the portfolio-choice and information-acquisition problem. The relation between costly information acquisition and the excess covariance of asset prices was studied in Veldkamp (2006). In Huang

[☆] The material in this paper was partially presented at none. This paper was recommended for publication in revised form by Associate Editor Valery Ugrinovskii under the direction of Editor Ian R. Petersen.

* Corresponding author.

E-mail addresses: 52194404014@stu.ecnu.edu.cn (Y. Zhang), zhuo.jin@mq.edu.au (Z. Jin), jqwei@stat.ecnu.edu.cn (J. Wei), gyin@uconn.edu (G. Yin).

and Liu (2007) where a dynamic portfolio-choice problem with static costly information choice was studied, the author assumed that the accuracy and frequency of information was chosen at one time only. However, in Gargano and Rossi (2018), an investor's attention was assumed time varying. Recently, a dynamic portfolio-choice problem with dynamic information acquisition was considered in Andrei and Hasler (2020). At each point in time, the investor optimally chose the quantity of information that she needed, which results in a dynamic trade-off problem between asset and attention allocation.

In this paper, we consider the time-inconsistent MV portfolio selection in a hidden Markov model with dynamic attention behavior. Following the concept of investor's attention to news in Andrei and Hasler (2020), we introduce the dynamic attention behavior viewed as an endogenous control into the hidden Markov model. We assume that the investor knows more accurate information about the current state of the market as her attention to news increases. Under this framework, the extended Hamilton–Jacobi–Bellman (HJB, for short) equation cannot be solved in closed-form. So our main objective is to find numerical approximation of the resulting control problem. To obtain the numerical solutions, we adopt the Markov chain approximation method developed in Kushner and Dupuis (2013), which has been used to find optimal controls in various complex stochastic systems in finance, insurance, and other fields, see Jin, Yang, and Yin (2013), Jin, Yin, and Wu (2013), Jin, Yin, and Zhu (2012), Song (2008), Song, Yin, and Zhang (2006), Tran and Yin (2016), among others. To the best of our knowledge, this paper is the first work to study extended HJB equation numerically by using the Markov chain approximation method, while the aforementioned references are all about solving the classical HJB equation. We build an iterative algorithm and obtain the convergence of the algorithm. Besides, numerical examples are provided to illustrate the results. The numerical results show that an investor with more wealth tends to acquire more information, which is consistent with those in Andrei and Hasler (2020).

The remainder of this paper is organized as follows. Section 2 introduces the model with the investor's attention to news. The Markov chain approximation method is considered in Section 3. In Section 4, we consider the approximation of optimal controls and establish the convergence of the algorithm. Section 5 gives one numerical example for illustration. Finally, Section 6 concludes the paper with further remarks.

2. Model

Let $T > 0$ be a fixed finite time horizon and $\mathbf{W}_1(\cdot) = (W_1^1(\cdot), \dots, W_1^d(\cdot))^T$ be a d -dimensional standard Brownian motion, where the symbol \top indicates the transpose.¹ Let $\alpha(t)$ be a continuous-time Markov chain independent of \mathbf{W}_1 valued in the finite set $\mathcal{M} = \{1, 2, \dots, m\}$ and denote by $\mathbf{Q} = (q^{ij})_{m \times m}$ the generator of $\alpha(t)$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space on which \mathbf{W}_1 , $\alpha(t)$, and all the random variables and processes in the rest of the paper are defined.

We consider a financial market consisting of one bond and d stocks within the time horizon $[0, T]$. The price of the risk-free bond S_0 satisfies

$$dS_0 = r(s, \alpha(s)) S_0 ds, \quad s \in (0, T],$$

where $S_0(0) = s_0 > 0$ and $r(\cdot, i) \geq 0$ is the risk-free return rate in state i for each $i = 1, 2, \dots, m$.

For $l = 1, 2, \dots, d$, and $s \in (0, T]$, the price of the l th stock S_l is given by

$$dS_l = \mu_l(s, \alpha(s)) S_l ds + \sum_{j=1}^d \sigma_{lj}(s, \alpha(s)) S_l dW_1^j,$$

where $S_l(0) = s_l > 0$, $\mu_l(\cdot, i)$ and $\sigma_{lj}(\cdot, i)$, $i = 1, 2, \dots, m$ are the expected return and volatility rate of the l th risky asset corresponding to state i , respectively.

In our framework, instead of having full information of the Markov chain $\alpha(t)$, the investor can only observe it in white noise. However, we allow the investor to have opportunity to actively learn about regime predictability.

Similar to Andrei and Hasler (2020), we assume that the investor may acquire $n(s)$ signals of equal precision $Z^j(s)$, $j = 1, \dots, n(s)$ at time s , and each signal represents information from market, which is given by

$$dZ^j = \zeta(\alpha(s)) ds + \sigma dB^j, \quad j = 1, \dots, n(s),$$

where B^j for $j = 1, \dots, n(s)$ are independent real-valued standard Brownian motions that are independent of \mathbf{W}_1 . Here, the function ζ is given by a vector $\zeta = (\zeta_1, \zeta_2, \dots, \zeta_m)'$, so that $\zeta(\alpha) = \langle \alpha, \zeta \rangle$ where $\langle \cdot, \cdot \rangle$ denotes the scalar product in \mathbb{R}^m . Then there is a standard real-valued Brownian motion B so that the aggregated signal $Z(s) := \sum_{j=1}^{n(s)} dZ^j(s)/n(s)$ acquired by the investor satisfies

$$dZ = \zeta(\alpha(s)) ds + \sigma / \sqrt{n(s)} dB. \quad (1)$$

Therefore the investor can choose the number of signals $n(s)$ to control the accuracy of the aggregated signal. If we set $\pi(s) = n(s)/\sigma^2$. Then clearly, the diffusion coefficient in (1) is $1/\sqrt{\pi(s)}$. To take this idea further, we shall not restrict $\pi(s)$ to take only discrete values when $n(s)$ is integer-valued, but consider it to take all real values in an arbitrary interval. This can be considered as a “fluid approximation”. It is more convenient for us to work with a real-valued control function than an integer-valued function.

Thus, we choose diffusion term as a control and consider the following signal process

$$dZ = \zeta(\alpha(s)) ds + 1/\sqrt{\pi} dB, \quad s \in (0, T],$$

with $Z(0) = 0$. Here, we assume that $\pi \in [\epsilon, M]$, where $\epsilon, M > 0$ are two arbitrary constants. The boundedness of π is consistent with Sims (2003), which argued that investors have limited information-processing capacity.

We further assume that the investor can get accurate signals at the cost of decreasing her current wealth, i.e., the investor faces a dynamic trade-off problem of asset and attention allocation. Denote by \bar{K} the total information cost and assume that $\bar{K} = K(\pi)X$, where the per-unit-of-wealth cost function $K(\pi)$ is increasing and convex in attention (Andrei & Hasler, 2020). Let u_l be the dollar amount invested in the l th stock. Then the wealth of the investor, denoted by X , evolves as

$$dX = b(s, X, \alpha(s), \mathbf{u}, \pi) ds + \sigma(s, X, \alpha(s), \mathbf{u}) d\mathbf{W}_1, \quad (2)$$

where $X(0) = x_0$, $\mathbf{u} := (u_1, \dots, u_d)^T$,

$$b(s, X, \alpha, \mathbf{u}, \pi) := r(s, \alpha) X + \theta^\top(s, \alpha) \mathbf{u} - \bar{K},$$

$$\sigma(s, X, \alpha, \mathbf{u}) := \mathbf{u}^\top (\sigma_{lj}(s, \alpha))_{1 \leq l, j \leq d}.$$

Here, $\theta(\cdot, i) := (\mu_1(\cdot, i) - r(\cdot, i), \dots, \mu_d(\cdot, i) - r(\cdot, i))^\top$, for $i = 1, 2, \dots, m$.

Next, let us recall some results about Wonham's filter. To proceed, we denote by I_E the indicator function of the event E and define $p^i(t) := I_{\{\alpha(t)=i\}}$, and

$$\varphi^i(t) := \mathbb{P}(\alpha(t) = i | Z(s), 0 \leq s \leq t).$$

¹ In the following, we shall suppress the suffix (\cdot) and the time variable for processes and functions whenever there is no confusion.

Following [Tran and Yin \(2016\)](#), we know that $\varphi^i(t)$, $i = 1, 2, \dots$, m is the probability conditioned on the observation $\sigma(Z(s), 0 \leq s \leq t)$. Since $\varphi^i(t) \geq 0$ and $\sum_{i=1}^m \varphi^i(t) = 1$, it is sufficient to work with $\varphi := (\varphi^1, \dots, \varphi^{m-1})^\top$. Denote by

$$S_{m-1} := \left\{ \varphi = (\varphi^1, \dots, \varphi^{m-1})^\top : \varphi^i \geq 0, \sum_{i=1}^{m-1} \varphi^i \leq 1 \right\},$$

and write $\varphi^m = 1 - \sum_{i=1}^{m-1} \varphi^i$ for $\varphi \in S_{m-1}$.

It was shown in [Wonham \(1964\)](#) that the posterior probability φ satisfies the following system of stochastic differential equation

$$d\varphi = \tilde{b}(\varphi) ds + \tilde{\sigma}(\varphi, \pi) d\hat{W}_2, \quad s \in (0, T],$$

where $\varphi(0) = \varphi_0 := (\varphi_0^1, \dots, \varphi_0^{m-1})^\top$ is the initial distribution of $\alpha(t)$,

$$\tilde{b}(\varphi) := \left(\sum_{j=1}^m q^{i1} \varphi^j, \dots, \sum_{j=1}^m q^{i(m-1)} \varphi^j \right)^\top,$$

$$\tilde{\sigma}(\varphi, \pi) := \left(\sqrt{\pi} \varphi^1 (\zeta(1) - \bar{\zeta}(\varphi)), \dots, \sqrt{\pi} \varphi^{m-1} (\zeta(m-1) - \bar{\zeta}(\varphi)) \right)^\top,$$

$\bar{\zeta}(\varphi) := \sum_{i=1}^m \zeta(i) \varphi^i$, and the innovation process \hat{W}_2 , independent of \mathbf{W}_1 , is defined as

$$\hat{W}_2(t) := \int_0^t \sqrt{\pi(s)} dZ(s) - \int_0^t \sqrt{\pi(s)} \bar{\zeta}(\varphi(s)) ds.$$

With $p^i(t)$, (2) can be written as

$$dX = \sum_{i=1}^m p^i [b(s, X, i, \mathbf{u}, \pi) ds + \sigma(s, X, i, \mathbf{u}) d\mathbf{W}_1].$$

By replacing the hidden state p^i by its estimate φ^i , we obtain an estimated process of the original process. For notational simplicity, we still denote by X the estimated process. Letting $\mathbf{Y} := (X, \varphi)^\top$, then we can get a completely observed system with the following compacted form

$$d\mathbf{Y} = \mathbf{f}(s, \mathbf{Y}, \mathbf{u}, \pi) ds + \Sigma(s, \mathbf{Y}, \mathbf{u}, \pi) d\hat{\mathbf{v}}, \quad (3)$$

where $\mathbf{Y}(0) = \mathbf{y}_0 := (x_0, \varphi_0)^\top$, $\hat{\mathbf{v}} := (\mathbf{W}_1, \hat{W}_2)^\top$,

$$\mathbf{f}(s, \mathbf{Y}, \mathbf{u}, \pi) := \left(\tilde{b}(s, \mathbf{Y}, \mathbf{u}, \pi), \tilde{\sigma}(\varphi, \pi) \right)^\top,$$

$$\Sigma(s, \mathbf{Y}, \mathbf{u}, \pi) := \begin{pmatrix} \tilde{\sigma}(s, \mathbf{Y}, \mathbf{u}) & 0 \\ \mathbf{0} & \tilde{\sigma}(\varphi, \pi) \end{pmatrix}.$$

Here, we define $\tilde{b}(s, \mathbf{Y}, \mathbf{u}, \pi) := \sum_{i=1}^m b(s, x, i, \mathbf{u}, \pi) \varphi^i$ and $\tilde{\sigma}(s, \mathbf{Y}, \mathbf{u}) := \sum_{i=1}^m \sigma(s, x, i, \mathbf{u}) \varphi^i$.

Let $\mathbb{F} := \{\mathcal{F}_s : s \in [0, T]\}$, where $\mathcal{F}_s = \{\mathbf{W}_1(\tilde{s}), Z(\tilde{s}), X(t) : t \leq \tilde{s} \leq s\}$. For $\mathbf{H} := \mathbb{R}, \mathbb{R}^d$, etc. and $0 \leq t \leq s \leq T$, define

$$L_{\mathbb{F}}^2(t, s; \mathbf{H}) := \{X : [t, s] \times \Omega \rightarrow \mathbf{H} \mid X \text{ is } \mathbb{F}\text{-adapted,}$$

$$\mathbb{E} \left[\int_t^s |X(v)|^2 dv \right] < \infty \},$$

$$L_{\mathbb{F}}^2(\Omega; C([t, s]; \mathbf{H})) := \{X : [t, s] \times \Omega \rightarrow \mathbf{H} \mid X \text{ is}$$

\mathbb{F} -adapted, has continuous paths, and

$$\mathbb{E} \left[\sup_{v \in [t, s]} |X(v)|^2 \right] < \infty \}.$$

In this paper, we restrict ourselves to closed-loop controls, that is at each t , $(\mathbf{u}(t), \pi(t))$ is given by $\mathbf{u}(t) = \tilde{\mathbf{u}}(t, \mathbf{Y}(t))$, $\pi(t) = \tilde{\pi}(t, \mathbf{Y}(t))$, where the maps $\tilde{\mathbf{u}} : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}_+^d$ and $\tilde{\pi} : [0, T] \times \mathbb{R}^m \rightarrow [\epsilon, M]$ are two Borel measurable functions. Here, we call the functions $\tilde{\mathbf{u}}$ and $\tilde{\pi}$ control laws. In the following, for notational simplicity, we still write \mathbf{u} and π for $\tilde{\mathbf{u}}$ and $\tilde{\pi}$, respectively.

Definition 1. Admissible control laws are maps \mathbf{u} and π satisfying the following conditions:

- (1) For each initial state $(t, \mathbf{y}) \in [0, T] \times \mathbb{R}^m$, (3) has a unique solution $\mathbf{Y} \in L_{\mathbb{F}}^2(\Omega; C([t, T]; \mathbb{R}^d))$.
- (2) The processes $\mathbf{u}(\cdot, \mathbf{Y}(\cdot)) \in L_{\mathbb{F}}^2(0, T; \mathbb{R}_+^d)$ and $\pi(\cdot, \mathbf{Y}(\cdot)) \in L_{\mathbb{F}}^2(0, T; [\epsilon, M])$.

We denote by \mathbf{U} and Π the class of admissible control laws \mathbf{u} and π , respectively.

For any initial state (t, \mathbf{y}) , the objective of the investor is to minimize the MV cost functional

$$J(t, \mathbf{y}; \mathbf{u}, \pi) = \text{Var}_t[X(T)] - \frac{\gamma}{2} \mathbb{E}_t[X(T)], \quad t \in [0, T], \quad (4)$$

where $\gamma > 0$ is a constant and $\mathbb{E}_t[\cdot] := \mathbb{E}[\cdot | \mathcal{F}_t]$. Letting $F(x) = x - \frac{\gamma}{2} x^2$ and $G(x) = \frac{\gamma}{2} x^2$, we can rewrite (4) as

$$J(t, \mathbf{y}; \mathbf{u}, \pi) = \mathbb{E}_t[F(X(T))] - G[\mathbb{E}_t(X(T))]. \quad (5)$$

In summary, we need to solve the following problem

Problem 1. At any time $t \in [0, T]$ with wealth and posterior probability $\mathbf{Y}(t) = (X(t), \varphi(t))$, specify admissible control laws $\mathbf{u} \in \mathbf{U}$ and $\pi \in \Pi$ that minimize (5).

It is well-known that [Problem 1](#) is time-inconsistent as the term $G[\mathbb{E}_t(X(T))]$ involves a non-linear function of the expectation. In this paper we aim to find the time-consistent equilibrium strategies for this problem. Following [Björk and Murgoci \(2010\)](#), we give the definition of equilibrium strategies and equilibrium value function.

Definition 2. Consider admissible control laws $\tilde{\mathbf{u}}$ and $\tilde{\pi}$. For arbitrary admissible control laws $\mathbf{u} \in \mathbf{U}$, $\pi \in \Pi$ and a fixed real number $\epsilon > 0$, define the control laws $(\mathbf{u}_\epsilon, \pi_\epsilon)$ by

$$(\mathbf{u}_\epsilon(s, \mathbf{y}), \pi_\epsilon(s, \mathbf{y})) = \begin{cases} (\mathbf{u}(s, \mathbf{y}), \pi(s, \mathbf{y})) & \text{if } t \leq s < t + \epsilon, \mathbf{y} \in \mathbb{R}^m, \\ (\tilde{\mathbf{u}}(s, \mathbf{y}), \tilde{\pi}(s, \mathbf{y})) & \text{if } t + \epsilon \leq s \leq T, \mathbf{y} \in \mathbb{R}^m. \end{cases}$$

If

$$\liminf_{\epsilon \rightarrow 0} \frac{J(t, \mathbf{y}; \tilde{\mathbf{u}}, \tilde{\pi}) - J(t, \mathbf{y}; \mathbf{u}_\epsilon, \pi_\epsilon)}{\epsilon} \geq 0$$

for arbitrary $\epsilon > 0$ and $(t, \mathbf{y}) \in [0, T] \times \mathbb{R}^m$, we say that $(\tilde{\mathbf{u}}, \tilde{\pi})$ is an equilibrium control law. The corresponding equilibrium value function V is given by $V(t, \mathbf{y}) := J(t, \mathbf{y}; \tilde{\mathbf{u}}, \tilde{\pi})$.

For any twice continuously differentiable function $\phi(\cdot, \cdot, \cdot) : \mathbb{R}_+ \times \mathbb{R} \times S_{m-1} \mapsto \mathbb{R}$, we define

$$\begin{aligned} \mathcal{L}^{\mathbf{u}, \pi} \phi(t, \mathbf{y}) &:= \frac{\partial \phi}{\partial t} + \frac{\partial \phi}{\partial x} \tilde{b}(t, \mathbf{y}, \mathbf{u}, \pi) \\ &+ \frac{1}{2} \frac{\partial^2 \phi}{\partial x^2} \tilde{\sigma}(t, \mathbf{y}, \mathbf{u}) \tilde{\sigma}^\top(t, \mathbf{y}, \mathbf{u}) + \sum_{i=1}^{m-1} \frac{\partial \phi}{\partial \varphi^i} \sum_{j=1}^m q^{ij} \varphi^j \\ &+ \frac{1}{2} \pi \sum_{i,k=1}^{m-1} \varphi^i \varphi^k (\zeta(i) - \bar{\zeta}(\varphi)) (\zeta(k) - \bar{\zeta}(\varphi)) \frac{\partial^2 \phi}{\partial \varphi^i \partial \varphi^k}. \end{aligned}$$

From the general theory of Markovian time-inconsistent stochastic control (see, e.g., [Björk and Murgoci \(2010\)](#)), we know that the equilibrium value function $V(t, \mathbf{y})$ satisfies the following system of extended HJB equation

$$\begin{aligned} \inf_{\mathbf{u}, \pi} \left\{ -\frac{\gamma}{2} \frac{\partial^2 g(t, \mathbf{y})}{\partial x^2} \tilde{\sigma}(t, \mathbf{y}, \mathbf{u}) \tilde{\sigma}^\top(t, \mathbf{y}, \mathbf{u}) + \mathcal{L}^{\mathbf{u}, \pi} V(t, \mathbf{y}) \right. \\ \left. - \frac{\gamma}{2} \pi \sum_{i,k=1}^{m-1} \varphi^i (\zeta(i) - \bar{\zeta}(\varphi)) \varphi^k (\zeta(k) - \bar{\zeta}(\varphi)) \frac{\partial^2 g(t, \mathbf{y})}{\partial \varphi^i \partial \varphi^k} \right\} \\ = 0, \end{aligned} \quad (6)$$

$$V(T, \mathbf{Y}(T)) = x, \quad \mathcal{L}^{\tilde{\mathbf{u}}, \tilde{\pi}} g(t, \mathbf{y}) = 0, \quad g(T, \mathbf{Y}(T)) = x,$$

where $\bar{\mathbf{u}}$ and $\bar{\pi}$ are control laws which realize the infimum in the first equation of (6).

Remark 1. Although we adopt the similar idea of dynamic attention behavior as in Andrei and Hasler (2020), we work under a hidden Markov model and use the Wonham filter to estimate the posterior probability. This is different from Andrei and Hasler (2020) where the author uses the Kalman filter. As we know, the Wonham filter is a non-linear filter, which makes it more difficult to handle. Besides, we study equilibrium strategies for a MV problem via extended HJB equation while Andrei and Hasler (2020) considers a classical utility maximization problem.

3. Discrete-time approximation scheme

In this section, we present a numerical algorithm to solve (6) using Markov chain approximation method. First, we introduce a pair of step sizes $h = (h_1, h_2)$. Here, $h_1 > 0$ is the discretization parameter for state variables, $h_2 > 0$ is the step size for time variable such that $N_{h_2} = T/h_2$ is an integer without loss of generality. Define $\mathbf{S}_{h_1} := \{(k^1 h_1, \dots, k^m h_1)^\top : k^i = 0, \pm 1, \dots, i = 1, \dots, m\}$ and let $\{\xi_n^h, n < \infty\}$ be a discrete-time controlled Markov chain with state space \mathbf{S}_{h_1} . Let $\mathbf{u}^h = (\mathbf{u}_0^h, \mathbf{u}_1^h, \dots)$ and $\pi^h = (\pi_0^h, \pi_1^h, \dots)$ denote the sequences of \mathbb{R}_+^d and $[\epsilon, M]$ valued random variables that are the control actions at time $0, 1, \dots$, respectively.

Denote by $\mathbf{P}^h((\mathbf{y}, \mathbf{z}) | \mathbf{r}, c)$ the probability that ξ transits from state \mathbf{y} at time nh_2 to state \mathbf{z} at time $(n+1)h_2$ conditioned on $\mathbf{u}_n^h = \mathbf{r}$ and $\pi_n^h = c$. Let \mathbf{U}^h and Π^h denote the collection of ordinary controls, which are determined by measurable function Λ_n^h , such that $(\mathbf{u}_n^h, \pi_n^h) = \Lambda_n^h(\xi_k^h, k \leq n, \mathbf{u}_k^h, \pi_k^h, k < n)$. We say that \mathbf{u}_n^h and π_n^h are admissible for the chain if \mathbf{u}_n^h and π_n^h are \mathbb{R}_+^d and $[\epsilon, M]$ valued random variables, respectively, and the Markov property still holds, namely,

$$\begin{aligned} & \mathbf{P}\{\xi_{n+1}^h = \mathbf{z} | \xi_n^h, \mathbf{u}_n^h, \pi_n^h, k \leq n\} \\ &= \mathbf{P}\{\xi_{n+1}^h = \mathbf{z} | \xi_n^h, \mathbf{u}_n^h, \pi_n^h\} := \mathbf{P}^h((\xi_n^h, \mathbf{z}) | \mathbf{u}_n^h, \pi_n^h). \end{aligned}$$

By using the Markov chain above, we can approximate the objective function defined in (5) by

$$J(t, \mathbf{y}; \mathbf{u}^h, \pi^h) = \mathbb{E}_t[F(\eta_{N_{h_2}}^h)] - G[\mathbb{E}_t(\eta_{N_{h_2}}^h)],$$

where η is the first component of ξ and $\eta_{N_{h_2}}^h$ is the terminal value of the discretized wealth process.

The sequence $\{\xi_n^h, n < \infty\}$ is said to be locally consistent (see, e.g., Kushner and Dupuis (2013)) w.r.t. (3), if it satisfies

$$\begin{aligned} \mathbb{E}_{\mathbf{y}, n}^{h, \mathbf{r}, c} \Delta \xi_n^h &= \mathbf{f}(t, \mathbf{y}, \mathbf{r}, c) h_2 + o(h_2), \\ \text{Var}_{\mathbf{y}, n}^{h, \mathbf{r}, c} \Delta \xi_n^h &= \Sigma(t, \mathbf{y}, \mathbf{r}, c) \Sigma(t, \mathbf{y}, \mathbf{r}, c)^\top h_2 + o(h_2), \\ \sup_n |\Delta \xi_n^h| &\rightarrow 0, \quad \text{as } h \rightarrow 0, \end{aligned}$$

where $\Delta \xi_n^h := \xi_{n+1}^h - \xi_n^h$, $\mathbb{E}_{\mathbf{y}, n}^{h, \mathbf{r}, c}$ and $\text{Var}_{\mathbf{y}, n}^{h, \mathbf{r}, c}$ denote the conditional expectation and variance given by $\{\xi_k^h, \mathbf{u}_k^h, \pi_k^h, k \leq n, \xi_n^h = \mathbf{y}, \mathbf{u}_n^h = \mathbf{r}, \pi_n^h = c\}$, respectively.

Now, we need to find transition probabilities such that the approximating Markov chain constructed above is locally consistent. To proceed, we first suppose that control space has a unique admissible pair control $(\hat{\mathbf{u}}, \hat{\pi})$ so that we can drop inf in (6). We discretize (6) by the following finite difference method using step sizes $h = (h_1, h_2)$.

Let $\mathbf{V} := (V, g)^\top$ and denote $\mathbf{V}^h(t, \mathbf{y})$ the solution to the finite difference equation, we have $\mathbf{V}(t, \mathbf{y}) \rightarrow \mathbf{V}^h(t, \mathbf{y})$.

For the derivative w.r.t. time variable, we use

$$\frac{\partial V^h(t, \mathbf{y})}{\partial t} \rightarrow \frac{V^h(t + h_2, \mathbf{y}) - V^h(t, \mathbf{y})}{h_2}.$$

For the first derivative w.r.t. x , we use one-side difference

$$\begin{aligned} & \frac{\partial V^h(t, \mathbf{y})}{\partial x} \\ & \rightarrow \begin{cases} \frac{V^h(t + h_2, x + h_1, \boldsymbol{\varphi}) - V^h(t + h_2, x, \boldsymbol{\varphi})}{h_1}, & \text{if } \bar{b}(t, \mathbf{y}, \hat{\mathbf{u}}, \hat{\pi}) \geq 0, \\ \frac{V^h(t + h_2, x, \boldsymbol{\varphi}) - V^h(t + h_2, x - h_1, \boldsymbol{\varphi})}{h_1}, & \text{if } \bar{b}(t, \mathbf{y}, \hat{\mathbf{u}}, \hat{\pi}) < 0. \end{cases} \end{aligned}$$

For the second derivative w.r.t. x , we have standard difference

$$\begin{aligned} & \frac{\partial^2 V^h(t, \mathbf{y})}{\partial x^2} \\ & \rightarrow \frac{\sum_{l=1}^2 V^h(t + h_2, x + (-1)^l h_1, \boldsymbol{\varphi}) - 2V^h(t + h_2, x, \boldsymbol{\varphi})}{h_1^2}. \end{aligned}$$

Similarly, we can approximate the first and second derivatives w.r.t. φ^i and all the partial derivatives of $g^h(t, \mathbf{y})$ (see also Tran and Yin (2016)). To simplify notations, we write $\mathbf{V}_n^h(\mathbf{y})$ for $\mathbf{V}^h(nh_2, \mathbf{y})$. Then, detailed calculation leads to the following iterative formula

$$\begin{aligned} \mathbf{V}_n^h(\mathbf{y}) &= (\bar{g}_{n+1}^h(\mathbf{y}), 0)^\top + \mathbf{V}_{n+1}^h(\mathbf{y}) p_1 \\ &+ \mathbf{V}_{n+1}^h(x + h_1, \boldsymbol{\varphi}) p_{2+} + \mathbf{V}_{n+1}^h(x - h_1, \boldsymbol{\varphi}) p_{2-} \\ &+ \sum_{i=1}^{m-1} [\mathbf{V}_{n+1}^h(x, \boldsymbol{\varphi} + h_1 \mathbf{e}_i) p_{3+}^i + \mathbf{V}_{n+1}^h(x, \boldsymbol{\varphi} - h_1 \mathbf{e}_i) p_{3-}^i] \\ &+ \sum_{l=1}^2 \sum_{k \neq i}^{m-1} [\mathbf{V}_{n+1}^h(x, \boldsymbol{\varphi} + (-1)^l h_1 (\mathbf{e}_i + \mathbf{e}_k)) p_{4+}^{i,k} \\ &+ \mathbf{V}_{n+1}^h(x, \boldsymbol{\varphi} + (-1)^l h_1 (\mathbf{e}_i - \mathbf{e}_k)) p_{4-}^{i,k}], \end{aligned}$$

where (suppressing variables $t, \mathbf{y}, \hat{\mathbf{u}}, \hat{\pi}$ of b and σ)

$$\begin{aligned} p_1 &:= \frac{h_2}{2h_1^2} \left[\hat{\pi} \sum_{i,k=1}^{m-1} \varphi^i \varphi^k |\zeta(i) - \bar{\zeta}(\boldsymbol{\varphi})| |\zeta(k) - \bar{\zeta}(\boldsymbol{\varphi})| \right. \\ &\quad \left. - 3\hat{\pi} \sum_{i=1}^{m-1} (\varphi^i |\zeta(i) - \bar{\zeta}(\boldsymbol{\varphi})|^2) \right] \\ &\quad - \frac{(|b| + \sum_{i=1}^{m-1} |\sum_{j=1}^m q^{ij} \varphi^j|) h_1 + \sigma \sigma^\top}{h_1^2} h_2 + 1, \\ p_{2\pm} &:= \frac{\sigma \sigma^\top + 2b^\pm h_1}{2h_1^2} h_2, \\ p_{3\pm}^i &:= \frac{h_2}{h_1^2} \left[\hat{\pi} \sum_{i=1}^{m-1} (\varphi^i |\zeta(i) - \bar{\zeta}(\boldsymbol{\varphi})|)^2 + (\sum_{j=1}^m q^{ij} \varphi^j)^\pm h_1 \right. \\ &\quad \left. - \frac{\hat{\pi} \varphi^i |\zeta(i) - \bar{\zeta}(\boldsymbol{\varphi})| \sum_{k=1}^{m-1} \varphi^k |\zeta(k) - \bar{\zeta}(\boldsymbol{\varphi})|}{2} \right], \\ p_{4\pm}^{i,k} &:= \frac{\hat{\pi} [\varphi^i (\zeta(i) - \bar{\zeta}(\boldsymbol{\varphi})) \varphi^k (\zeta(k) - \bar{\zeta}(\boldsymbol{\varphi}))]^\pm}{4h_1^2} h_2. \end{aligned} \tag{7}$$

Here, x^+ and x^- are the positive and negative parts of the real number x , respectively, and $\bar{g}_{n+1}^h(\mathbf{y})$ is given by

$$\begin{aligned} & \bar{g}_{n+1}^h(\mathbf{y}) \\ &= \gamma g_{n+1}^h(\mathbf{y}) \left(1 - p_1 - \frac{\sum_{i=1}^{m-1} |\sum_{j=1}^m q^{ij} \varphi^j| + |b|}{h_1} h_2 \right) \\ &+ \gamma \sum_{l=1}^2 \left\{ g_{n+1}^h(x + (-1)^l h_1, \boldsymbol{\varphi}) \frac{b^+ h_2 - p_{2+} h_1}{h_1} \right. \end{aligned}$$

$$\begin{aligned}
& + \sum_{i=1}^{m-1} g_{n+1}^h(x, \varphi + (-1)^i h_1 \mathbf{e}_i) \frac{(\sum_{j=1}^m q^{ij} \varphi^j)^+ h_2 - p_{3+}^i h_1}{h_1} \\
& - \sum_{k \neq i}^{m-1} \left[\mathbf{V}_{n+1}^h(x, \varphi + (-1)^i h_1 (\mathbf{e}_i + \mathbf{e}_k)) p_{4+}^{i,k} \right. \\
& \quad \left. + \mathbf{V}_{n+1}^h(x, \varphi + (-1)^i h_1 (\mathbf{e}_i - \mathbf{e}_k)) p_{4-}^{i,k} \right].
\end{aligned}$$

By choosing proper h , we can assume that the coefficient p_1 is in $[0, 1]$, which can be explained as the transition probability from state \mathbf{y} at time nh_2 to state \mathbf{y} at time $(n+1)h_2$ given $(\hat{\mathbf{u}}, \hat{\pi})$. Similarly, we can explain the transition probabilities $p_{2\pm}, p_{3\pm},$ and $p_{4\pm}^{i,k}$. Similar to Yang et al. (2015), we can approximate $V(t, \mathbf{y})$ by using

$$V(t, \mathbf{y}) = \inf_{\mathbf{u}^h \in \mathcal{U}^h, \pi^h \in \Pi^h} J(t, \mathbf{y}; \mathbf{u}^h, \pi^h).$$

It follows from Kushner and Dupuis (2013) that the Markov chain $\{\xi_n^h, n < \infty\}$ with transition probabilities defined in (7) is indeed locally consistent with (3).

4. Approximation of optimal controls

4.1. Relaxed controls

Following Kushner and Dupuis (2013), let us first recall the definition of relaxed controls. Let $\mathcal{B}(\mathbf{U})$ and $\mathcal{B}(\mathbf{U} \times [t, T])$ denote the Borel σ -algebras of \mathbf{U} and $\mathbf{U} \times [t, T]$, respectively. An admissible relaxed control or simply a relaxed control m is a measure on $\mathcal{B}(\mathbf{U} \times [t, T])$ such that $m(\mathbf{U} \times [t, s]) = s - t$ for all $s \in [t, T]$. Similarly we can define \tilde{m} a measure on $\mathcal{B}(\Pi \times [t, T])$ such that $\tilde{m}(\Pi \times [t, s]) = s - t$. Given relaxed controls m and \tilde{m} , there are derivatives m_s and \tilde{m}_s which are measures on $\mathcal{B}(\mathbf{U})$ and $\mathcal{B}(\Pi)$, respectively, such that $m(d\mathbf{r}, ds) = m_s(d\mathbf{r})ds$ and $\tilde{m}(d\mathbf{c}, ds) = \tilde{m}_s(d\mathbf{c})ds$, respectively.

To prove the convergence of the algorithm, we use relaxed control representation (Kushner & Dupuis, 2013). Let $\mathbf{M} = (M_1, \dots, M_d)^\top$ and $\tilde{\mathbf{M}}$ be a vector-valued measure and a scalar-valued measure, respectively. To proceed, we impose the following conditions.

(A1) \mathbf{M} is square integrable and continuous, each component is orthogonal, and the pairs $(M_i, M_j)_{i \neq j}$ and $\tilde{\mathbf{M}}$ are strongly orthogonal. $\tilde{\mathbf{M}}$ is also square integrable, continuous and orthogonal (see, e.g., Yang et al. (2015)).

Under Assumption (A1), there exist measure-valued random processes m^i and \tilde{m} such that the quadratic variation processes satisfy, for $i = 1, 2, \dots, d$,

$$\begin{aligned}
\langle M_i(A, \cdot), M_j(B, \cdot) \rangle(s) &= \delta_{ij} m^i(A \cap B, s), \\
\langle \tilde{M}(\tilde{A}, \cdot), \tilde{M}(\tilde{B}, \cdot) \rangle(s) &= \tilde{m}(\tilde{A} \cap \tilde{B}, s),
\end{aligned}$$

where $A, B \in \mathcal{B}(\mathbf{U})$, $\tilde{A}, \tilde{B} \in \mathcal{B}(\Pi)$ and δ_{ij} is Kronecker delta.

(A2) m^i does not depend on i , and $m(\mathbf{U}, s) = s$.

Under Assumption (A1) and (A2), there are measures \mathbf{M} and $\tilde{\mathbf{M}}$ with quadratic variation processes $m\mathbf{1}$ and \tilde{m} , respectively, where $\mathbf{1}$ is the unit vector.

With the relaxed controls representation, the operator of the controlled diffusion is given by

$$\begin{aligned}
& \mathcal{L}^{m, \tilde{m}} \phi(t, \mathbf{y}) \\
& := \frac{\partial \phi}{\partial t} + \int \int \frac{\partial \phi}{\partial \mathbf{x}} b(t, \mathbf{y}, \mathbf{r}, c) m_s(d\mathbf{r}) \tilde{m}_s(d\mathbf{c}) \\
& \quad + \frac{1}{2} \int \frac{\partial^2 \phi}{\partial \mathbf{x}^2} \sigma(t, \mathbf{y}, \mathbf{r}) \sigma^\top(t, \mathbf{y}, \mathbf{r}) m_s(d\mathbf{r}) \\
& \quad + \sum_{i=1}^{m-1} \frac{\partial \phi}{\partial \varphi^i} \sum_{j=1}^m q^{ij} \varphi^j + \frac{1}{2} \int \sum_{i,k=1}^{m-1} \varphi^i (\zeta(i) - \bar{\zeta}(\varphi)) \\
& \quad \times \varphi^k (\zeta(k) - \bar{\zeta}(\varphi)) \frac{\partial^2 \phi}{\partial \varphi^i \partial \varphi^k} c \tilde{m}_s(d\mathbf{c}) \\
& = \int \int \mathcal{L}^{\mathbf{u}, \pi} \phi(s, \mathbf{y}) m_s(d\mathbf{r}) \tilde{m}_s(d\mathbf{c}).
\end{aligned}$$

Besides, there are measures m, \tilde{m} satisfying Assumptions (A1) and (A2) such that for each bounded and smooth function $\tilde{f}(\cdot, \cdot)$,

$$\tilde{f}(s, \mathbf{Y}) - \tilde{f}(t, \mathbf{y}) - \int \int \int \mathcal{L}^{\mathbf{u}, \pi} \tilde{f}(z, \mathbf{Y}(z)) m_z(d\mathbf{r}) \tilde{m}_z(d\mathbf{c}) dz$$

is an $\tilde{\mathbb{F}}$ martingale, where $\tilde{\mathbb{F}} = \sigma\{\mathbf{Y}(z), m_z, \tilde{m}_z, t \leq z \leq s\}$.

Then $(\mathbf{Y}, m, \tilde{m})$ solves the martingale problem with operator $\mathcal{L}^{\mathbf{u}, \pi}$ and we represent our control system as

$$\begin{aligned}
\mathbf{Y}(s) &= \mathbf{y} + \int_t^s \int_{\mathbf{U}} \int_{\Pi} \mathbf{f}(\mathbf{Y}(z), \mathbf{r}, c) m_z(d\mathbf{r}) \tilde{m}_z(d\mathbf{c}) dz \\
& \quad + \int_t^s \int_{\mathbf{U}} \int_{\Pi} \Sigma(\mathbf{Y}(z), \mathbf{r}, c) \mathbf{M}(d\mathbf{r}, d\mathbf{z}) \tilde{\mathbf{M}}(d\mathbf{c}, d\mathbf{z}).
\end{aligned} \tag{8}$$

We say that (\mathbf{M}, m) and $(\tilde{\mathbf{M}}, \tilde{m})$ are admissible relaxed controls for (8) if Assumption (A1) and (A2) hold and $\langle \mathbf{M} \rangle = m\mathbf{1}$, $\langle \tilde{\mathbf{M}} \rangle = \tilde{m}$. To proceed, we pose another two conditions.

(A3) $\mathbf{f}(\cdot, \cdot, \cdot)$ and $\Sigma(\cdot, \cdot, \cdot)$ are continuous; $\mathbf{f}(\cdot, \mathbf{r}, c)$ and $\Sigma(\cdot, \mathbf{r}, c)$ are Lipschitz continuous uniformly in \mathbf{r}, c and bounded.

(A4) Let $\mathbf{a} := \Sigma \Sigma^\top$, there exist a positive constant C and a identity matrix \mathbf{I} such that $\mathbf{a} - C\mathbf{I}$ is positive definite and $\mathbf{a}_{ii} - \sum_{j:j \neq i} |\mathbf{a}_{ij}| \geq 0$, for each $i = 1, \dots, m$.

4.2. Approximation of relaxed controls and measures

Let E_n^h denote the conditional expectation given $\{\xi_k^h, \mathbf{u}_k^h, \pi_k^h, k \leq n\}$ and define $\mathbf{R}_n^h = \Delta \xi_n^h - E_n^h \Delta \xi_n^h$. By local consistency, we have

$$\xi_{n+1}^h = \xi_n^h + \mathbf{f}(\xi_n^h, \mathbf{u}_n^h, \pi_n^h) h_2 + \mathbf{R}_n^h,$$

with

$$\begin{aligned}
\text{Cov}_n^h \mathbf{R}_n^h &= \mathbf{a}(\xi_n^h, \mathbf{u}_n^h, \pi_n^h) \\
&= \Sigma(\xi_n^h, \mathbf{u}_n^h, \pi_n^h) \Sigma^\top(\xi_n^h, \mathbf{u}_n^h, \pi_n^h) h_2 + o(h).
\end{aligned}$$

As shown in Tran and Yin (2016), by Assumption (A4), we can decompose

$$\mathbf{a}(\xi_n^h, \mathbf{u}_n^h, \pi_n^h) = \mathbf{T}_n^h (\mathbf{D}_n^h)^2 (\mathbf{T}_n^h)^\top,$$

where $\mathbf{D}_n^h = \text{diag}(d^1, \dots, d^m)$ and \mathbf{T}_n^h is an orthogonal matrix. Then we can represent the increment of Brownian motion as $\Delta \mathbf{W}_n^h = (\mathbf{D}_n^h)^{-1} (\mathbf{T}_n^h)^\top \mathbf{R}_n^h$, where $\mathbf{R}_n^h = \Sigma(\xi_n^h, \mathbf{u}_n^h, \pi_n^h) \Delta \mathbf{W}_n^h + \mathbf{e}_n^h$. Thus we have

$$\xi_{n+1}^h = \xi_n^h + \mathbf{f}(\xi_n^h, \mathbf{u}_n^h, \pi_n^h) h_2 + \Sigma(\xi_n^h, \mathbf{u}_n^h, \pi_n^h) \Delta \mathbf{W}_n^h + \mathbf{e}_n^h.$$

To focus on the control part, let $\{\mathbf{L}_l^h, l \leq k_h\}$ and $\{C_l^h, \tilde{l} \leq \tilde{k}_h\}$ be a finite partition of \mathbf{U} and \mathcal{I} , respectively, such that $\mathbf{L}_l^h \rightarrow \mathbf{0}$ and $C_l^h \rightarrow 0$, as $h \rightarrow 0$. Let $\mathbf{r}_l \in \mathbf{L}_l^h$ and $c_l \in C_l^h$, we can define the random variable

$$\Delta \mathbf{W}_{l,\tilde{l},n}^h = \Delta \mathbf{W}_{n,l}^h I_{\{\mathbf{u}_n^h = \mathbf{r}_l, \pi_n^h = c_l\}} + \Delta \mathbf{W}_{n,l,\tilde{l}}^h I_{\{\mathbf{u}_n^h \neq \mathbf{r}_l, \pi_n^h \neq c_l\}}.$$

Then we have

$$\begin{aligned} \xi_{n+1}^h &= \xi_n^h + \mathbf{f}(\xi_n^h, \mathbf{u}_n^h, \pi_n^h) h_2 \\ &+ \sum_{l,\tilde{l}} \Sigma(\xi_n^h, \mathbf{u}_n^h, \pi_n^h) I_{\{\mathbf{u}_n^h = \mathbf{r}_l, \pi_n^h = c_l\}} \Delta \mathbf{W}_{l,\tilde{l},n}^h + \mathbf{e}_n^h, \\ m_n^h(\mathbf{r}_l) &= I_{\{\mathbf{u}_n^h = \mathbf{r}_l\}}, \quad \tilde{m}_n^h(c_l) = I_{\{\pi_n^h = c_l\}}. \end{aligned} \quad (9)$$

To approximate the continuous-time following processes $(\mathbf{Y}, m, \mathbf{M}, \tilde{m}, \tilde{M})$, for $s \in [nh_2, (n+1)h_2]$, we define the piecewise constant interpolations as

$$\begin{aligned} \Psi^h(s) &= \Psi_n^h, \text{ for } \Psi = \xi, \mathbf{u}, \pi, \mathbf{e}, m, \tilde{m}, \quad z^{h_2}(s) = n, \\ \bar{\alpha}^h(s) &= \sum_{i=1}^m g(i) \varphi_n^{h,i}, \quad \mathbf{W}_{l,\tilde{l}}^h(s) = \sum_{k=0}^{z^{h_2}(s)-1} \Delta \mathbf{W}_{l,\tilde{l},k}^h. \end{aligned} \quad (10)$$

Recall that m_n^h and \tilde{m}_n^h are a pair admissible relaxed controls if $m_n^h(\mathbf{U}) = 1, \tilde{m}_n^h(\mathcal{I}) = 1$ and

$$\begin{aligned} &\mathbb{P}\{\xi_{n+1}^h = \mathbf{z} | \xi_i^h, m_i^h, \tilde{m}_i^h, i \leq n\} \\ &= \int \mathbb{P}^h(\xi_n^h, \mathbf{z} | \mathbf{r}, c) m_n^h(d\mathbf{r}) \tilde{m}_n^h(dc). \end{aligned}$$

For $\mathbf{r}_l \in \mathbf{L}_l^h, c_l \in C_l^h, \{\mathbf{M}(\mathbf{L}_l^h, \cdot), l \leq k_h\}$ and $\{\tilde{M}(C_l^h, \cdot), \tilde{l} \leq \tilde{k}_h\}$ are orthogonal continuous martingales with $\langle \mathbf{M}(\mathbf{L}_l^h, \cdot) \rangle = m(\mathbf{L}_l^h, \cdot) \mathbf{1}$, and $\langle \tilde{M}(C_l^h, \cdot) \rangle = \tilde{m}(C_l^h, \cdot)$. There are $d+1$ dimensional standard Brownian motions $\mathbf{W}_{l,\tilde{l}}^h(\cdot), l \leq k_h, \tilde{l} \leq \tilde{k}_h$ such that

$$\mathbf{M}(\mathbf{L}_l^h, s) \tilde{M}(C_l^h, s) = \int_t^s m_z^{1/2}(\mathbf{L}_l^h) \tilde{m}_z^{1/2}(C_l^h) d\mathbf{W}_{l,\tilde{l}}^h(z).$$

Let \mathbf{M}^h and m^h be the restrictions of the measures of \mathbf{M} and m , respectively, on the set $\{\mathbf{L}_l^h, l \leq k_h\}$, and let \tilde{M}^h and \tilde{m}^h be the restrictions of the measures of \tilde{M} and \tilde{m} , respectively, on the set $\{C_l^h, \tilde{l} \leq \tilde{k}_h\}$. Similar to the method used in Kushner (1990, Theorem 8.1, pp.1024–1025), we have the following lemma.

Lemma 3. Under Assumption (A1)–(A4), it holds that

$$(\xi^h, m^h, \mathbf{M}^h, \tilde{m}^h, \tilde{M}^h) \Rightarrow (\mathbf{Y}, m, \mathbf{M}, \tilde{m}, \tilde{M}),$$

where the notation \Rightarrow denotes weak convergence, and $J(t, \mathbf{y}; m^h, \tilde{m}^h) \rightarrow J(t, \mathbf{y}; m, \tilde{m})$.

Lemma 3 shows we can approximate the variables $(\mathbf{Y}, m, \mathbf{M}, \tilde{m}, \tilde{M})$ by using $(\xi^h, m^h, \mathbf{M}^h, \tilde{m}^h, \tilde{M}^h)$ satisfying

$$\begin{aligned} \xi^h(s) &= \mathbf{y} + \int_t^s \sum_{l,\tilde{l}} \mathbf{f}(\xi^h(z), \mathbf{r}_l, c_l) m_z(\mathbf{L}_l^h) \tilde{m}_z(C_l^h) dz \\ &+ \int_t^s \sum_{l,\tilde{l}} \Sigma(\xi^h(z), \mathbf{r}_l, c_l) m_z^{1/2}(\mathbf{L}_l^h) \tilde{m}_z^{1/2}(C_l^h) d\mathbf{W}_{l,\tilde{l}}^h(z) \\ &+ \mathbf{e}^h(s). \end{aligned} \quad (11)$$

Let \mathbb{F}^h represent the σ -algebra that is generated by

$$\begin{aligned} &\{\xi^h(z), m_z^h, \tilde{m}_z^h, \mathbf{M}^h(z), \tilde{M}^h(z), \\ &\mathbf{W}_{l,\tilde{l}}^h(z), 1 \leq l \leq k_h, 1 \leq \tilde{l} \leq \tilde{k}_h, t \leq z \leq s\}. \end{aligned}$$

Denote Γ^h and $\tilde{\Gamma}^h$ the sets of admissible relaxed controls m^h and \tilde{m}^h w.r.t. the set $\{\xi^h, \mathbf{W}_{l,\tilde{l}}^h, l \leq k_h, \tilde{l} \leq \tilde{k}_h\}$, respectively such that m_s^h and \tilde{m}_s^h are fixed probability measures in the interval $[nh_2, (n+1)h_2]$. We then rewrite the value function as

$$V^h(t, \mathbf{y}) = \inf_{m^h \in \Gamma^h, \tilde{m}^h \in \tilde{\Gamma}^h} J^h(t, \mathbf{y}; m^h, \tilde{m}^h), \quad (12)$$

and $J^h(t, \mathbf{y}; m^h, \tilde{m}^h) = \mathbb{E}_t[F(\eta^h(T))] - G[\mathbb{E}_t(\eta^h(T))]$.

Next, we use the weak convergence methods (see, e.g., Kushner and Dupuis (2013)) to obtain the convergence of the algorithm. Let $(\xi^h, m^h, \mathbf{M}^h, \tilde{m}^h, \tilde{M}^h)$ be a solution of (11), where \mathbf{M}^h and \tilde{M}^h are measures with respect to the filtration \mathbb{F}^h , with quadratic variation processes $m^h \mathbf{1}$ and \tilde{m}^h , respectively. Then we have the following two theorems, whose proofs are presented in Appendix A and Appendix B of Zhang, Jin, Wei, and Yin (2022).

Theorem 4. Under Assumption (A1)–(A4), let the approximating chain $\{\xi_n^h, n < \infty\}$ be constructed with transition probabilities defined in (7), let $\{\mathbf{u}_n^h, n < \infty\}$ and $\{\pi_n^h, n < \infty\}$ be sequences of admissible controls, ξ^h be the continuous time interpolations defined by (10), m^h be the relaxed control representation of \mathbf{u}^h (continuous time interpolation of \mathbf{u}_n^h), and \tilde{m}^h be the relaxed control representation of π^h (continuous time interpolation of π_n^h). Then $H^h := \{\xi^h, m^h, \mathbf{M}^h, \tilde{m}^h, \tilde{M}^h\}$ is tight, which has a weakly convergent subsequence with the limit $H := \{\mathbf{Y}, m, \mathbf{M}, \tilde{m}, \tilde{M}\}$.

Theorem 5. Under Assumption (A1)–(A4), let $V(t, \mathbf{y})$ and $V^h(t, \mathbf{y})$ be value functions defined in (5) and (12), respectively. Then $V^h(t, \mathbf{y}) \rightarrow V(t, \mathbf{y})$, as $h \rightarrow 0$.

5. A numerical example

In this section, we provide an example with $m = 2$ to demonstrate our results. Here, we set $h_1 = 0.2, h_2 = 0.001, \gamma = 0.5, N = 2000$ and $\pi \in [0.001, 2]$. We consider the same financial market as Yang et al. (2015).

The per-unit-of-wealth information-cost function is specified in quadratic form, i.e., $\bar{K}(\pi) = k\pi^2$, where $k > 0$ is the information-cost parameter (see Andrei and Hasler (2020)). Fix $t = 1$. Let $\phi := \varphi^1(1) = 0.2$, the value function $V(1)$, the ratio of risky investment to wealth $w(1)$ and the control $\pi(1)$ with different values of k are shown in Figs. (a)–(c), respectively. The relationship between the value function $V(1)$ and (ϕ, x) is shown in Fig. (d).

Fig. (a) shows the value of value function is larger when k is smaller. This is because smaller k implies lower information cost, and thus the investor can obtain better MV utility. It follows from Fig. (b) that an investor intends to invest in risky assets as wealth increases. This may be due to the fact that the investor with less wealth intends to invest in risk-free assets for risk-averse purposes. Given fixed wealth x , larger k implies higher information cost, and thus the investor intends to invest in risky assets, which makes larger w . Fig. (c) shows that as wealth increases, the investor pays more attention to the signals with small k . Because of the specific assumption of the information cost function in this section, the price of information is relatively high with larger k when investor's wealth reaches a certain level. This will lead the investor to reduce investment in the information market more quickly.

6. Concluding remarks

This paper has investigated time-consistent equilibrium strategies for MV portfolio selection under a hidden Markov model. We adopted the idea of the dynamic attention behavior in Andrei and

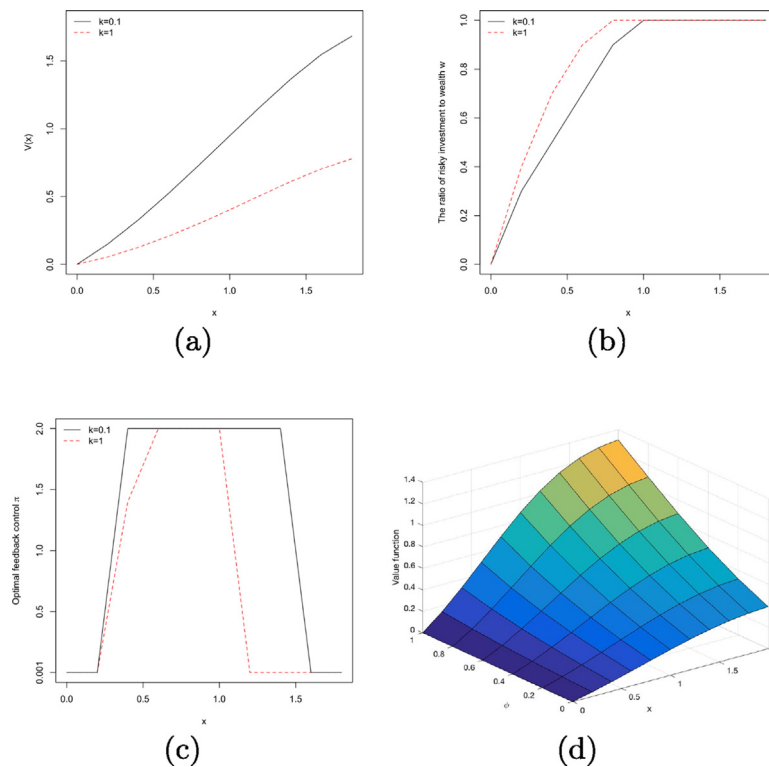


Fig. 1. Optimal controls and value function.

Hasler (2020) to introduce investor's attention to news on the hidden Markov model. That is, we considered an investor who can, at each time, improve the accuracy of acquired information at a cost. Under this framework, we derived an extended HJB equation, for which we used Markov chain approximation to obtain numerical solutions. We constructed an iterative algorithm, proved its convergence, and provided some numerical results. We also gave some explanations for our numerical results.

Acknowledgments

The authors are grateful to the anonymous reviewers and the editors whose valuable comments and suggestions have helped improving the paper and clarifying the presentation. This work was supported by the Fundamental Research Funds for the Central Universities (2020ECNU-HWFW003), the 111 Project (B14019) and the National Natural Science Foundation of China (12071146, 11571113, 11601157, 11601320), Research Grants Council of the Hong Kong Special Administrative Region (project No. 17304921). The research of G. Yin was supported in part by the National Science Foundation of the United States under grant DMS-2114649.

References

- Andrei, D., & Hasler, M. (2020). Dynamic attention behavior under return predictability. *Management Science*, 66(7), 2906–2928.
- Björk, T., & Murgoci, A. (2010). A general theory of Markovian time-inconsistent stochastic control problems. *Preprint SSRN*, 55.
- Detemple, J., & Kihlstrom, R. (1987). Acquisition d'information dans un modèle intertemporel en temps continu. *L'Actualité Économique*, 63(2–3), 118–137.
- Ekeland, I., & Lazrak, A. (2006). Being serious about non-commitment: Subgame perfect equilibrium in continuous time. *ArXiv preprint arXiv:math/0604264*.
- Ekeland, I., & Pirvu, T. A. (2008). Investment and consumption without commitment. *Mathematics and Financial Economics*, 2(1), 57–86.
- Elliott, R. J., Siu, T. K., & Badescu, A. (2010). On mean-variance portfolio selection under a hidden Markovian regime-switching model. *Economic Modelling*, 27(3), 678–686.
- Gargano, A., & Rossi, A. G. (2018). Does it pay to pay attention? *Review of Financial Studies*, 31(12), 4595–4649.
- Hu, Y., Jin, H., & Zhou, X. Y. (2012). Time-inconsistent stochastic linear-quadratic control. *SIAM Journal on Control and Optimization*, 50(3), 1548–1572.
- Huang, L., & Liu, H. (2007). Rational inattention and portfolio selection. *The Journal of Finance*, 62(4), 1999–2040.
- 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., Yin, G., & Wu, F. (2013). Optimal reinsurance strategies in regime-switching jump diffusion models: Stochastic differential game formulation and numerical methods. *Insurance: Mathematics & Economics*, 53(3), 733–746.
- Jin, Z., Yin, G., & Zhu, C. (2012). Numerical solutions of optimal risk control and dividend optimization policies under a generalized singular control formulation. *Automatica*, 48(8), 1489–1501.
- Kushner, H. (1990). Numerical methods for stochastic control problems in continuous time. *SIAM Journal on Control and Optimization*, 28(5), 999–1048.
- Kushner, H., & Dupuis, P. G. (2013). *Numerical methods for stochastic control problems in continuous time* (Vol. 24). Springer Science & Business Media.
- Li, D., & Ng, W. (2000). Optimal dynamic portfolio selection: Multiperiod mean-variance formulation. *Mathematical Finance*, 10(3), 387–406.
- Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77–91.
- Sims, C. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50(3), 665–690.
- Song, Q. (2008). Convergence of Markov chain approximation on generalized HJB equation and its applications. *Automatica*, 44(3), 761–766.
- Song, Q., Yin, G., & Zhang, Z. (2006). Numerical methods for controlled regime-switching diffusions and regime-switching jump diffusions. *Automatica*, 42(7), 1147–1157.
- Strotz, R. (1955). Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies*, 23(3), 165–180.
- Tran, K., & Yin, G. (2016). Numerical methods for optimal harvesting strategies in random environments under partial observations. *Automatica*, 70, 74–85.
- Veldkamp, L. (2006). Information markets and the comovement of asset prices. *Review of Economic Studies*, 73(3), 823–845.
- Wang, T., Jin, Z., & Wei, J. (2019). Mean-variance portfolio selection under a non-Markovian regime-switching model: Time-consistent solutions. *SIAM Journal on Control and Optimization*, 57(5), 3249–3271.

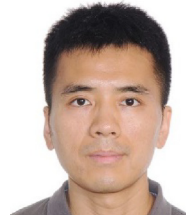
- Wei, J., & Wang, T. (2017). Time-consistent mean-variance asset-liability management with random coefficients. *Insurance: Mathematics & Economics*, 77, 84–96.
- Wonham, W. (1964). Some applications of stochastic differential equations to optimal nonlinear filtering. *Journal of the Society for Industrial and Applied Mathematics, Series A: Control*, 2(3), 347–369.
- Xia, Y. (2001). Learning about predictability: The effects of parameter uncertainty on dynamic asset allocation. *The Journal of Finance*, 56(1), 205–246.
- Yang, Z., Yin, G., & Zhang, Q. (2015). Mean-variance type controls involving a hidden Markov chain: Models and numerical approximation. *IMA Journal of Mathematical Control and Information*, 32(4), 867–888.
- Zhang, Y., Jin, Z., Wei, J., & Yin, G. (2022). Mean-variance portfolio selection with dynamic attention behavior in a hidden Markov model. ArXiv preprint arXiv:2205.08743.
- Zhou, X. Y., & Li, D. (2000). Continuous-time mean-variance portfolio selection: A stochastic LQ framework. *Applied Mathematics and Optimization*, 42(1), 19–33.
- Zhou, X. Y., & Yin, G. (2003). Markowitz's mean-variance portfolio selection with regime switching: A continuous-time model. *SIAM Journal on Control and Optimization*, 42(4), 1466–1482.



Yu Zhang received the B.Ec. degree in Financial Statistics from Jiangxi University of Finance and Economics, and the M.Sc. degree in Statistics from East China Normal University, in 2017 and 2019, respectively. He is currently pursuing a Ph.D. in Actuarial Science at East China Normal University. His current research interests include stochastic optimal control and mathematical finance.



Zhuo Jin received the B.S. degree in mathematics from the Huazhong University of Science and Technology in 2005 and Ph.D. in Mathematics from Wayne State University in 2011. He joined the department of actuarial studies and business analytics at Macquarie University as a full professor in 2022 after he worked as a lecturer, senior lecturer, and associate professor at The University of Melbourne for ten years. His research interests include stochastic optimal control, actuarial science, and mathematical finance. He is an Associate in the Society of Actuaries (ASA).



Jiaqin Wei received his B.S. degree in Mathematical Statistics in 2006 and Ph.D. in Actuarial Science in 2011 both from East China Normal University. He is now a full professor in the School of Statistics at East China Normal University. Before he joined East China Normal University as a research professor in 2016, he worked as a postdoctoral research fellow in the Department of Applied Finance and Actuarial Studies at Macquarie University for 3 years. His research fields include actuarial science and mathematical finance.



George Yin received the Ph.D. degree in applied mathematics from Brown University in 1987, joined Wayne State University in 1987, became University Distinguished Professor in 2017, and moved to the University of Connecticut in 2020. His research interests include stochastic systems theory and applications. Dr. Yin was the Chair of the SIAM Activity Group on Control and Systems Theory, and served on the Board of Directors of the American Automatic Control Council. He is the Editor-in-Chief of *SIAM J. Control Optim.* and is (or was) on the editorial board of many other journals including *Automatica* (2005–2011) and *IEEE Trans. Automatic Control* (1994–1998). He is a Fellow of IFAC, a Fellow of IEEE, and a Fellow of SIAM.