Markowitz's Mean-Variance Portfolio Selection with Regime Switching: From Discrete-time Models to Their Continuous-time Limits

G. Yin, *Fellow, IEEE,* and Xun Yu Zhou, *Senior Member, IEEE*

***Abstract*— We study a discrete-time version of Markowitz's mean-variance portfolio selection problem where the market parameters depend on the market mode (regime) that jumps among a finite number of states. The random regime switching is delineated by a finite-state Markov chain, based on which a discrete-time Markov modulated portfolio selection model is presented. Such models either arise from multiperiod portfolio selections or result from numerical solution of continuous-time problems. The natural connections between discrete-time models and their continuous-time counterpart are revealed. Since the Markov chain frequently has a large state space, to reduce the complexity, an aggregated process with smaller state space is introduced and the underlying portfolio selection is formulated as a two-time-scale problem. We prove that the process of interest yields a switching diffusion limit using weak convergence methods. Next, based on the optimal control of the limit process obtained from our recent work, we devise portfolio selection strategies for the original problem and demonstrate their asymptotic optimality.**

***Index Terms*— Markowitz's mean-variance portfolio selection, discrete-time model, Markov chain, switching diffusion, linear-quadratic problem, singular perturbation.**

I. INTRODUCTION

MARKOWITZ'S Nobel-prize winning mean-variance portfolio selection model (for a single period) [16], [17] provides a foundation of modern finance theory; it has inspired numerous extensions and applications. The Markowitz model aims to maximize the terminal wealth, in the mean time to minimize the risk using the variance as a criterion, which enables investors to seek highest return upon specifying their acceptable risk level.

There have been continuing efforts in extending portfolio selection from the static single period model to dynamic multi-period or continuous-time models. However, the research works on dynamic portfolio selections have been dominated by those of maximizing expected utility functions of the terminal wealth, which is in spirit different from the original Markowitz's model. For example, the multi-period utility models were investigated in Mossin [19], Samuelson [21], Hakansson [10], Elton and Gruber [7], Francis [8], Grauer and Hakansson [9], and Pliska [20], among many others. As for the continuous-time case, the famous Merton paper [18] studied a utility maximization problem with market factors modelled as a diffusion process (rather than as a Markov chain). Along another line, the mean-variance hedging problem was investigated by Duffie and Richardson [6] and Schweizer [22], where an optimal dynamic strategy was sought to hedge contingent claims in an imperfect market. Optimal hedging policies [6], [22] were obtained primarily based on the so-called projection theorem.

Very recently, using the stochastic linear-quadratic (LQ) theory developed in [4], [28], Zhou and Li [31] introduced a stochastic LQ control framework to study the continuous-time version of the Markowitz's problem. Within this framework, they derived closed-form efficient policies (in the Markowitz sense) along with an explicit expression of the efficient frontier.

In the aforementioned references, for continuous-time formulations of mean-variance problems, stochastic differential equations and geometric Brownian motion models were used. Although such models have been used in a wide variety of situations, they have certain limitations since all the key parameters, including the interest rate and the stock appreciation/volatility rates, are assumed to be insensitive to the (very likely) drastic changes in the market. Typically, the underlying market may have many "modes" or "regimes" that switch among themselves from time to time. The market mode could reflect the state of the underlying economy, the general mood of investors in the market, and so on. For example, the market can be roughly divided as "bullish" and "bearish", while the market parameters can be quite different in the two modes. One could certainly introduce more intermediate states between the two extremes. A system, commonly referred to as the regime switching model, can be formulated as a stochastic differential equation whose coefficients are modulated by a continuous-time Markov chain. Such a model has been employed in the literature to discuss options; see [1], [3], [5]. Moreover, an investment-consumption model with regime switching was studied in [29]; an optimal stock selling rule for a Markov-modulated Black-Scholes model was derived in [30]; a stochastic approximation approach for the liquidation problem could be found in [25]. In [32], we treated the continuous-time version of Markowitz's mean-variance portfolio selection with regime switching and derived the efficient portfolio and efficient frontier explicitly.

Motivated by the recent developments of mean-variance

G. Yin is with the Department of Mathematics, Wayne State University, Detroit, MI 48202, <gyin@math.wayne.edu>. The research of this author was supported in part by the National Science Foundation under grant DMS-9877090.

X.Y. Zhou is with the Department of Systems Engineering and Engineering Management, the Chinese University of Hong Kong, Shatin, Hong Kong, <xyzhou@se.cuhk.edu.hk>. The research of this author was supported in part by the RGC Earmarked Grants CUHK 4435/99E and CUHK 4175/00E.

portfolio selection and Markov-modulated geometric Brownian motion formulation, we develop a class of discrete-time mean-variance portfolio selection models and reveal their relationship with the continuous-time counterparts in this paper. The discrete-time case is as important as the continuous-time one. First, frequently, one needs to deal with multi-period, discrete-time Markowitz's portfolio selection problems directly; see Li and Ng [13] for a recent account on the topic, in which efficient strategies were derived together with the efficient frontier. In addition, to simulate a continuous-time model, one often has to use a discretization technique leading to a discrete-time problem formulation. In this paper, however, one of the main features of the problem to be tackled is that all the market coefficients are modulated by a discrete-time Markov chain that has a finite state space.

Owing to the presence of long-term and short-term investors, the movements of a capital market can be divided into primary movement and secondary movement naturally leading to two-time scales. Besides, with various economic factors such as trends of the market, interest rates, and business cycles being taken into consideration, the state space of the Markov chain, representing the totality of the possible market modes, is often large. If we simply treated each possible mode as an individual one distinct from all others, the size of the problem would be huge. A straightforward implementation of numerical schemes may deem to be infeasible due to the curse of dimensionality. It is thus crucial to find a viable alternative. To reduce the complexity, we observe that the transition rates among different states could be quite different. In fact, there is certain hierarchy (in terms of the magnitude of the transition rates) involved. Therefore, it is possible to lump many states at a similar hierarchical level together to form a big "state." With this aggregated Markov chain the size of the state space is substantially reduced. Now, to highlight the different rates of changes, we introduce a small parameter $\varepsilon > 0$ into the transition matrix, resulting in a singular perturbation formulation. Based on the recent progress on two-time-scale Markov chains (see [23], [26]), we establish the natural connection between the discrete-time problem and its continuous-time limit. Under simple conditions, we show that suitably interpolated processes converge weakly to their limits leading to a continuous-time mean-variance portfolio selection problem with regime switching. The limit mean-variance portfolio selection problem has an optimal solution [32] that can be obtained in a very simple way under appropriate conditions. Using that solution, we design policies that are asymptotically optimal. Our findings indicate that in lieu of examining the more complex original problem, we could use the much simplified limit problem as a guide to obtain portfolio selection policies that are nearly as good as the optimal one from a practical concern. The advantage is that the complexity is much reduced.

We remark that although the specific mean-variance portfolio selection is treated in this paper, the formulation and techniques can be generally employed as well in the so-called hybrid control problems that are modulated by a Markov chain for many other applications.

The rest of the work is arranged as follows. Section 2 begins with the formulation of the discrete-time mean-variance portfolio selection problem. Section 3 proceeds with some preliminary results concerning two-time-scale Markov chains and introduces an auxiliary process. Section 4 is devoted to weak convergence analysis, in which we establish the natural connection between the discrete-time and continuous-time models aiming at reducing the complexity of the underlying systems. Section 5 constructs policies that are based on optimal control of the limit problem and derives asymptotic optimal strategy via the constructed controls. Section 6 extends the results by allowing the Markov chain to be nonhomogeneous and/or including transient states. Section 7 concludes paper with additional remarks. To make the paper more accessible, a couple of long and technical proofs are placed in an appendix.

II. FORMULATION

Suppose that T is a fixed positive real number and that $\varepsilon > 0$ is a small parameter. Working with discrete time k, we consider $0 \le k \le \lfloor T/\varepsilon \rfloor$, where $\lfloor v \rfloor$ denotes the integer part of a real number v. For ease of presentation, in what follows, we suppress the floor function notation $\lfloor \cdot \rfloor$ whenever there is no confusion. All the random variables/processes in this paper are defined on a given complete probability space $(\Omega, \mathcal{F}, P)$.

Consider a market model as follows. Let $\mathcal{M} = \{1, 2, \cdots, m\}$ denote the collection of different market modes. Let α_k^ε, for $0 \le k \le T/\varepsilon$, be a discrete-time Markov chain, which is parameterized by ε, with the state space $\mathcal{M}$. Suppose that there are $d+1$ assets in the underlying market. One of which is the bond and the rest of them are the stock holdings. Use $S_k^{\varepsilon,0}$ to denote the price of the bond, and $S_k^{\varepsilon,\imath}$, $\imath = 1, \ldots, d$, to denote the prices of the stocks at time k, where respectively. (Note that here k does not represent the calendar time; it is the iteration time at which the systems dynamics are updated. The calendar time is εk.) Corresponding to a market mode $\ell \in \mathcal{M}$, let $r(\cdot, \ell)$ be the interest rate, and $b^\imath(\cdot, \ell)$ and $\sigma^{\imath\jmath}(\cdot, \ell)$ the appreciation rates and volatility rates of the stocks, respectively, where $r(\cdot,\cdot)$, $b^\imath(\cdot,\cdot)$, $\sigma^{\imath\jmath}(\cdot,\cdot) : \mathbb{R} \times \mathcal{M} \mapsto \mathbb{R}$, for $\imath, \jmath = 1, \ldots, d$, are given functions. For each $(t, \ell) \in \mathbb{R} \times \mathcal{M}$, denote

$$c^\imath(t,\ell) = \left[b^\imath(t,\ell) - \frac{1}{2} \sum_{\jmath=1}^{d} (\sigma^{\imath\jmath}(t,\ell))^2 \right], \imath = 1, \ldots, d. \quad (1)$$

Then under the so-called *multiplicative model*, the asset prices $S_k^{\varepsilon,\imath}$ satisfy the following system of equations:

$$\begin{aligned} &S_{k+1}^{\varepsilon,0} = S_k^{\varepsilon,0} + \varepsilon r(\varepsilon k, \alpha_k^\varepsilon) S_k^{\varepsilon,0}, \quad S_0^{\varepsilon,0} = S_0^0 > 0, \\ &S_{k+1}^{\varepsilon,\imath} = S_k^{\varepsilon,\imath} \exp\left(\varepsilon c^\imath(\varepsilon k, \alpha_k^\varepsilon) + \sqrt{\varepsilon} \sum_{\jmath=1}^{d} \sigma^{\imath\jmath}(\varepsilon k, \alpha_k^\varepsilon) \xi_k^\jmath \right), \\ &S_0^{\varepsilon,\imath} = S_0^\imath > 0, \quad \imath = 1, \ldots, d, \end{aligned} \quad (2)$$

where $\{\xi_k^\imath\}$, $\imath = 1, \ldots, d$, are sequences of independent and identically distributed random variables. Note that the multiplicative model adopted above is analogous to the geometric Brownian motion model in continuous time, which ensures the non-negativity of the stock prices.

Remark 2.1: Another frequently used model for discrete-time stock prices is the following:

$$\begin{aligned} &S^{\varepsilon,0}_{k+1} = S^{\varepsilon,0}_k + \varepsilon r(\varepsilon k, \alpha^\varepsilon_k) S^{\varepsilon,0}_k, \quad S^{\varepsilon,0}_0 = S^0 > 0, \\ &S^{\varepsilon,\imath}_{k+1} = S^{\varepsilon,\imath}_k + \varepsilon b^\imath(\varepsilon k, \alpha^\varepsilon_k) S^{\varepsilon,\imath}_k + \sqrt{\varepsilon} \sum_{\jmath=1}^d \sigma^{\imath\jmath}(\varepsilon k, \alpha^\varepsilon_k)\xi^\jmath_k S^{\varepsilon,\imath}_k, \\ &S^{\varepsilon,\imath}_0 = S^\imath > 0, \quad \imath = 1, \ldots, d. \end{aligned} \tag{3}$$

This model does not necessarily produce non-negative stock prices. However, as demonstrated in [15, p. 311] the two models are essentially the same in terms of approximating the stock price process.

Suppose that at time k, an investor with an initial endowment $x^\varepsilon_0 = x_0$ holds $N^\imath(\varepsilon k)$ shares of the $\imath$th asset, $\imath = 0, \ldots, d$, during the time interval $[k, k+1)$. Thus his or her total wealth at time k is $x^\varepsilon_k = \sum_{i=0}^d N^\imath(\varepsilon k) S^{\varepsilon,\imath}_k$ for $0 \le k \le T/\varepsilon$. Now for a self-financed portfolio (i.e., no infusion or withdrawal of funds during the indicated time interval), the difference in total wealth between two consecutive times is purely due to the change in the prices of the stocks (see, e.g., Karatzas and Shreve [11, p.6, Eq. (2.2)]). Therefore, we have

$$x^\varepsilon_{k+1} - x^\varepsilon_k = \sum_{\imath=0}^d N^\imath(\varepsilon k)(S^{\varepsilon,\imath}_{k+1} - S^{\varepsilon,\imath}_k) \quad \text{and} \tag{4}$$

$$x^\varepsilon_k = \sum_{\imath=0}^d u^{\varepsilon,\imath}_k, \; u^{\varepsilon,\imath}_k = N^\imath(\varepsilon k) S^{\varepsilon,\imath}_k \text{ for } \imath = 0, \ldots, d. \tag{5}$$

Denote by $\mathcal{F}^\varepsilon_k$, the σ-algebra generated by $\{\alpha^\varepsilon_{k_1}, \xi_{k_1} : 0 \le k_1 < k\}$, where $\xi_k = (\xi^1_k, \ldots, \xi^d_k)'$. A portfolio $u^\varepsilon_\cdot = \{u^\varepsilon_k = (u^{\varepsilon,1}_k, u^{\varepsilon,2}_k, \cdots, u^{\varepsilon,d}_k), \; k = 0, 1, \cdots, T/\varepsilon\}$ is admissible if u^ε_k is $\mathcal{F}^\varepsilon_k$-measurable for each $0 \le k \le T/\varepsilon$. Note that we do not include $u^{\varepsilon,0}_k$, the amount of money allocated to the bond, in defining the portfolio. This is because by (4), we have

$$\begin{aligned} x^\varepsilon_{k+1} &= x^\varepsilon_k + \varepsilon r(\varepsilon k, \alpha^\varepsilon_k) N^0(\varepsilon k) S^{\varepsilon,0}_k \\ &\quad + \sum_{\imath=1}^d N^\imath(\varepsilon k)(S^{\varepsilon,\imath}_{k+1} - S^{\varepsilon,\imath}_k) \\ &= x^\varepsilon_k + \varepsilon r(\varepsilon k, \alpha^\varepsilon_k)(x^\varepsilon_k - \sum_{\imath=1}^d u^{\varepsilon,\imath}_k) \\ &\quad + \sum_{\imath=1}^d N^\imath(\varepsilon k)(S^{\varepsilon,\imath}_{k+1} - S^{\varepsilon,\imath}_k). \end{aligned}$$

Therefore,

$$\begin{aligned} x^\varepsilon_{k+1} &= [1 + \varepsilon r(\varepsilon k, \alpha^\varepsilon_k)] x^\varepsilon_k - \varepsilon r(\varepsilon k, \alpha^\varepsilon_k) \sum_{\imath=1}^d u^{\varepsilon,\imath}_k \\ &\quad + \sum_{\imath=1}^d N^\imath(\varepsilon k)(S^{\varepsilon,\imath}_{k+1} - S^{\varepsilon,\imath}_k). \end{aligned}$$

Hence the wealth process $x^\varepsilon_\cdot = \{x^\varepsilon_k, \; k = 0, 1, \cdots, T/\varepsilon\}$ is completely determined by the allocation of wealth among the stocks (excluding the bond). We call $(x^\varepsilon_\cdot, u^\varepsilon_\cdot)$ an admissible wealth-portfolio pair, and denote the class of admissible wealth-portfolio pairs by $\mathcal{A}^\varepsilon$.

The objective of the discrete-time mean-variance portfolio selection problem is to find an admissible portfolio $(x^\varepsilon_\cdot, u^\varepsilon_\cdot) \in \mathcal{A}^\varepsilon$, for an initial wealth $x^\varepsilon_0 = x_0$ and an initial market mode $\alpha^\varepsilon_0 = \ell_0$, such that the terminal wealth is $Ex^\varepsilon_{T/\varepsilon} = z$ for a given $z \in \mathbb{R}$, and the risk in terms of the variance of the terminal wealth, $E[x^\varepsilon_{T/\varepsilon} - z]^2$, is minimized. That is,

$$\begin{aligned} &\text{Minimize } J^\varepsilon(x_0, \ell_0, u^\varepsilon_\cdot) = E[x^\varepsilon_{T/\varepsilon} - z]^2 \\ &\text{subject to: } x^\varepsilon_0 = x_0, \; \alpha^\varepsilon_0 = \ell_0, \; Ex^\varepsilon_{T/\varepsilon} = z \\ &\text{and } \; (x^\varepsilon_\cdot, u^\varepsilon_\cdot) \in \mathcal{A}^\varepsilon. \end{aligned} \tag{6}$$

III. Preliminary Results

Our effort in what follows is to obtain a limit problem of (6) as $\varepsilon \to 0$. The idea comes from aggregation of the Markovian states to reduce the complexity of the underlying system via a hierarchical approach. To proceed, we make the following assumptions. The first one is on the transition matrix of the Markov chain, the second one concerns about the interest rate of the bond, and the stock appreciation and volatility rates, and the last one is a condition on the exogenous noise $\{\xi^\imath_k\}$.

(A1) The transition matrix of the discrete-time Markov chain α^ε_k is given by

$$P^\varepsilon = P + \varepsilon Q, \tag{7}$$

$$P = \text{diag}(P^1, \ldots, P^l) \tag{8}$$

such that P^i for $i = 1, \ldots, l$ are transition matrices, and Q is a generator (i.e., $q^{\ell\ell_1} \ge 0$ for $\ell \ne \ell_1$ and $\sum_{\ell_1 \in \mathcal{M}} q^{\ell\ell_1} = 0$ for each $\ell \in \mathcal{M}$). Moreover, P^i, $i = 1, \ldots, l$ are irreducible and aperiodic.

(A2) For each $\ell \in \mathcal{M}$, $\imath, \jmath = 1, \ldots, d$, $r(\cdot, \ell)$, $b^\imath(\cdot, \ell)$, $\sigma^{\imath\jmath}(\cdot, \ell)$ are real-valued continuous functions defined on $[0, T]$.

(A3) For each $\imath = 1, \ldots, d$, $\{\xi^\imath_k\}$ is a sequence of independent and identically distributed (i.i.d.) random variables that are independent of α^ε_k and that have mean 0 and variance 1. Moreover, for $\imath \ne \jmath$, $\xi^\imath_k$ and $\xi^\jmath_k$ are independent.

Remark 3.1: The rationale for (A1) is that among the transition rates of the Markovian states, some of them vary rapidly and others change slowly. Decomposing the state space in accordance with these transition rates, we can write the state space as

$$\mathcal{M} = \mathcal{M}_1 \cup \mathcal{M}_2 \cdots \cup \mathcal{M}_l, \tag{9}$$

where $\mathcal{M}_i = \{\zeta_{i1}, \ldots, \zeta_{im_i}\}$. Since in a finite-state Markov chain, there is at least one recurrent state (not all states can be transient), and there is no null-recurrent state, the irreducibility implies that there are l ergodic classes of states. Each of the transition matrix P^i is responsible for the rapid transitions, whereas the generator Q governs the slow transitions from one ergodic class to another. The transition matrix P^ε given in (7) with P specified in (8) has the so-called nearly completely decomposable structure. It arises from many discrete-time control and optimization problems, or from discretization of a continuous-time problem; see [24] and the references therein. On the other hand, Condition (A3) indicates that the sequences $\{\xi^\imath_k\}$ are the so-called white noise. Under this condition, the Donsker's invariance principle [2, p.137] implies that $\sqrt{\varepsilon} \sum_{k=0}^{t/\varepsilon - 1} \xi^\imath_k$ converges weakly to a Brownian motion as $\varepsilon \to 0$. For definition, discussion, and basic results on weak convergence, we refer the reader to [2], [12]. In fact, correlated noise may also be dealt with. What is essential is a functional central limit theorem as mentioned above holds for the scaled sequence. Sufficient conditions guaranteeing the convergence, for example, for ϕ-mixing processes are available. However, assumption (A3) allows us to simplify much of the discussion in what follows.

Due to the various considerations, the state space $\mathcal{M}$ is large. That is, m is a large number. In solving the portfolio selection problem, we aggregate the states in each of the recurrent class $\mathcal{M}_i$ into one state, resulting in an aggregated process having only l states. If $l \ll m$, the possible number of configurations is substantially smaller for the aggregated process. In practice, we typically merge all the similar states into one "big" state. For example, we can put all the states together where the market is generally up (respectively, down) to form a "bullish" (respectively, "bearish") mode of the market. Our effort in what follows is to show that suitably interpolated processes of the prices converge to limit processes that are switching diffusions, and to establish the connection between the discrete-time and continuous-time portfolio selection problems. Using the optimal strategy of the limit problem, which was obtained explicitly in [32], enables us to construct an asymptotically optimal strategy for the original problem (6).

A. Singularly Perturbed Markov Chains

To carry out the analysis, define an aggregated discrete-time process $\overline{\alpha}^\varepsilon_k$ by $\overline{\alpha}^\varepsilon_k = i$ if $\alpha^\varepsilon_k \in \mathcal{M}_i, \quad i = 1, 2, \cdots, l$. Next define the interpolated processes of the stock/bond price processes, the aggregated process, and admissible wealth-portfolio pairs as: For $t \in [\varepsilon k, \varepsilon k + \varepsilon)$,

$$\begin{array}{ll} S^{\varepsilon,\imath}(t) = S^{\varepsilon,\imath}_k, \ \imath = 0, \ldots, d, & \overline{\alpha}^\varepsilon(t) = \overline{\alpha}^\varepsilon_k \\ u^\varepsilon(t) = u^\varepsilon_k, & x^\varepsilon(t) = x^\varepsilon_k. \end{array} \quad (10)$$

That is, they are all piecewise constant on the interval of length ε. We will obtain the weak convergence of the interpolated processes. Before proceeding further, we present some preliminary results.

Lemma 3.2: Assume condition (A1). Then the following assertions hold:

(i) Denote by ν^i the stationary distribution corresponding to the transition matrix P^i for each $i = 1, \ldots, l$. Then $p^\varepsilon_k = (P(\alpha^\varepsilon_k = \zeta_{11}), \ldots, P(\alpha^\varepsilon_k = \zeta_{lm_l})) \in \mathbb{R}^{1\times m}$ satisfies

$$p^\varepsilon_k = \theta(t)\mathrm{diag}(\nu^1, \ldots, \nu^l) + O(\varepsilon + \lambda^k), \quad (11)$$

for some $0 < \lambda < 1$, where $\theta(t) = (\theta_1(t), \ldots, \theta_l(t)) \in \mathbb{R}^{1\times l}$ (with $t = \varepsilon k$) satisfies

$$\frac{d\theta(t)}{dt} = \theta(t)\overline{Q}, \quad \theta_i(0) = x^i_0 \mathbb{1}_{m_i},$$
$$\overline{Q} = \mathrm{diag}(\nu^1, \ldots, \nu^l)Q\widetilde{\mathbb{1}}, \ \widetilde{\mathbb{1}} = \mathrm{diag}(\mathbb{1}_{m_1}, \ldots, \mathbb{1}_{m_l}). \quad (12)$$

Here $\mathbb{1}_\ell$ denotes an ℓ-dimensional column vector with all entries being 1.

(ii) For $k \le T/\varepsilon$, the k-step transition probability matrix $(P^\varepsilon)^k$ satisfies

$$(P^\varepsilon)^k = \Phi(t) + O\left(\varepsilon + \lambda^k\right), \quad (13)$$

where $\Phi(t) = \widetilde{\mathbb{1}}\Theta(t)\mathrm{diag}(\nu^1, \ldots, \nu^l)$ with $\Theta(t)$ satisfying

$$\frac{d\Theta(t)}{dt} = \Theta(t)\overline{Q}, \quad \Theta(0) = I. \quad (14)$$

(iii) As $\varepsilon \to 0$, $\overline{\alpha}^\varepsilon(\cdot)$ converges weakly to $\overline{\alpha}(\cdot)$, a continuous-time Markov chain with state space $\overline{\mathcal{M}} = \{1, \ldots, l\}$ and generator $\overline{Q}$ given by (12). Moreover, for the occupation measures defined by

$$o^\varepsilon_{k,ij} = \varepsilon \sum_{k_1=0}^{k} [I_{\{\alpha^\varepsilon_{k_1} = \zeta_{ij}\}} - \nu^i_j I_{\{\alpha^\varepsilon_k \in \mathcal{M}_i\}}],$$

with $i = 1, \ldots, l$, $j = 1, \ldots, m_i$, and $0 \le k \le T/\varepsilon$, the following mean square estimates hold

$$\sup_{0\le k\le T/\varepsilon} E|o^\varepsilon_{k,ij}|^2 = O(\varepsilon). \quad (15)$$

Proof. The proofs of (i) and (ii) are in [24], and that of (iii) can be found in [27]. □

B. An Auxiliary Process

The model (2) is a multiplicative one. For the subsequent analysis, it is more convenient to work with an auxiliary *additive model*. To do so we retain the definition of $S^{\varepsilon,0}_k$, and define auxiliary processes $Y^{\varepsilon,\imath}_k$ for $\imath = 1, \ldots, d$ as: $Y^{\varepsilon,\imath}_0 = 0$, and

$$Y^{\varepsilon,\imath}_{k+1} = Y^{\varepsilon,\imath}_k + \varepsilon c^\imath(\varepsilon k, \alpha^\varepsilon_k) + \sqrt{\varepsilon}\sum_{\jmath=1}^{d} \sigma^{\imath\jmath}(\varepsilon k, \alpha^\varepsilon_k)\xi^\jmath_k, \quad (16)$$

where $c^\imath(\cdot,\cdot)$ is given by (1). In terms of the processes $\{Y^{\varepsilon,\imath}_k\}$, the price for the $\imath$th stock can be written as

$$S^{\varepsilon,\imath}_k = S^\imath_0 \exp(Y^{\varepsilon,\imath}_k), \quad S^{\varepsilon,\imath}_0 = S^\imath_0 > 0. \quad (17)$$

Similar to the interpolation $S^{\varepsilon,\imath}(\cdot)$, define the interpolated processes

$$Y^{\varepsilon,\imath}(t) = Y^{\varepsilon,\imath}_k \ \text{ for } \ t \in [k\varepsilon, k\varepsilon + \varepsilon), \ \imath = 1, \ldots, d. \quad (18)$$

It is easily seen that $S^{\varepsilon,\imath}(t) = S^\imath_0 \exp(Y^{\varepsilon,\imath}(t))$. That is, $S^{\varepsilon,\imath}(\cdot)$ is related to the interpolation of $Y^{\varepsilon,\imath}_k$ through the exponential function.

To obtain the convergence of $S^{\varepsilon,\imath}(\cdot)$, we utilize the auxiliary process $Y^{\varepsilon,\imath}(\cdot)$. We first show that $\{Y^{\varepsilon,\imath}(\cdot)\}$ converges weakly to certain limit process, in which the appreciation and volatility rates are averaged out with respect to the stationary measures of the corresponding Markov chains. Then by using a well-known continuous mapping theorem [2, Theorem 5.2, p.31], we obtain the desired result.

IV. Weak Convergence

This section is devoted to the weak convergence of interpolated processes $S^{\varepsilon,0}(\cdot)$ and $Y^{\varepsilon,\imath}(\cdot)$. We first verify the tightness of the underlying sequences. Then we characterize the limit by showing that they are solutions of certain martingale problems with appropriate generators. In fact, we work with the processes $(S^{\varepsilon,0}(\cdot), \overline{\alpha}^\varepsilon(\cdot))$ and $(Y^{\varepsilon,\imath}(\cdot), \overline{\alpha}^\varepsilon(\cdot))$ for $\imath = 1, \ldots, d$, respectively.

A. Tightness of $\{S^{\varepsilon,0}(\cdot),\overline{\alpha}^{\varepsilon}(\cdot)\}$ and $\{Y^{\varepsilon,\imath}(\cdot),\overline{\alpha}^{\varepsilon}(\cdot)\}$

Theorem 4.1: Under (A1)–(A3), $\{S^{\varepsilon,0}(\cdot),\overline{\alpha}^{\varepsilon}(\cdot)\}$ and $\{Y^{\varepsilon,\imath}(\cdot),\overline{\alpha}^{\varepsilon}(\cdot)\}$ for $\imath=1,\dots,d$ are tight on $D[0,T]$, where $D[0,T]$ is the space of functions that are right continuous, have left limits, endowed with the Skorohod topology.

Proof. The proof is in the appendix. □

Since $\{S^{\varepsilon,0}(\cdot),\overline{\alpha}^{\varepsilon}(\cdot)\}$ and $\{Y^{\varepsilon,\imath}(\cdot),\overline{\alpha}^{\varepsilon}(\cdot)\}$ ($\imath=1,\dots,d$) are tight, Prohorov's theorem ([2, p.37]) allows us to extract weakly convergent subsequences. Select such convergent subsequences (still indexed by ε for notational simplicity) with limits $(S^0(\cdot),\overline{\alpha}(\cdot))$ and $(Y^\imath(\cdot),\overline{\alpha}(\cdot))$, respectively. By virtue of the Skorohod representation ([12, p.29]), we may assume without loss of generality that $(S^{\varepsilon,0}(\cdot),\overline{\alpha}^{\varepsilon}(\cdot))$ and $(Y^{\varepsilon,\imath}(\cdot),\overline{\alpha}^{\varepsilon}(\cdot))$ converge to $(S^0(\cdot),\overline{\alpha}(\cdot))$ and $(Y^\imath(\cdot),\overline{\alpha}(\cdot))$, respectively, with probability one (w.p.1), and that they converge uniformly on any compact time interval.

B. Weak Convergence

We proceed to characterize the limit processes. Denote the $d\times d$ matrix $(\sigma^{\imath\jmath}(t,i))$ by $\Sigma(t,i)$. Define

$$\begin{aligned}\overline{r}(t,i)&=\sum_{j=1}^{m_i}\nu_j^i r(t,\zeta_{ij}),\\ \overline{b}^\imath(t,i)&=\sum_{j=1}^{m_i}\nu_j^i b^\imath(t,\zeta_{ij}),\\ \overline{c}^\imath(t,i)&=\sum_{j=1}^{m_i}\nu_j^i c^\imath(t,\zeta_{ij}),\end{aligned}\tag{19}$$

where ν_j^i denotes the jth component of the stationary distribution ν^i corresponding to P^i. Let $\overline{\Sigma}(t,\alpha)=(\overline{\sigma}^{\imath\jmath}(t,\alpha))$ be such that

$$\overline{\Sigma}(t,i)\overline{\Sigma}'(t,i)=\sum_{j=1}^{m_i}\nu_j^i\Sigma(t,\zeta_{ij})\Sigma'(t,\zeta_{ij}),\ i\in\overline{\mathcal{M}}=\{1,\dots,l\},\tag{20}$$

where for a vector or a matrix B, B' denotes its transpose. Concentrating on the processes $(S^{\varepsilon,0}(\cdot),\overline{\alpha}^{\varepsilon}(\cdot))$ and $(Y^{\varepsilon,\imath}(\cdot),\overline{\alpha}^{\varepsilon}(\cdot))$, we wish to show that the limit processes are solutions of

$$\begin{aligned}&\frac{dS^0(t)}{dt}=\overline{r}(t,\overline{\alpha}(t))S^0(t),\ S^0(0)=S_0^0,\\ &dY^\imath(t)=\overline{c}^\imath(t,\overline{\alpha}(t))dt+\sum_{\jmath=1}^{d}\overline{\sigma}^{\imath\jmath}(t,\overline{\alpha}(t))dw^\jmath(t),\\ &Y^\imath(0)=0,\qquad \imath=1,\dots,d,\end{aligned}\tag{21}$$

respectively, where $w^\imath(\cdot)$ for $\imath=1,\dots,d$ are independent, scalar, standard Brownian motions. Equivalently, it suffices to show that for each $\imath=0,\dots,d$, $(S^\imath(\cdot),\overline{\alpha}(\cdot))$ is a solution of the martingale problem with the operator $(\partial/\partial t)+\mathcal{L}^\imath$ and

$$\begin{aligned}\mathcal{L}^0f(t,y,i)&=f_y(t,y,i)\overline{r}(t,i)y,\\ \mathcal{L}^\imath f(t,y,i)&=f_y(t,y,i)\overline{c}^\imath(t,i)+\frac{f_{yy}(t,y,i)}{2}[\overline{\Sigma}(t,i)\overline{\Sigma}'(t,i)]^{\imath\imath}\\ &\quad+\overline{Q}f(t,y,\cdot)(i),\ 1\le\imath\le d,\end{aligned}\tag{22}$$

where for each i, $f(\cdot,\cdot,i)$ is a suitable real-valued function defined on $\mathbb{R}\times\mathbb{R}$, $\overline{Q}=(\overline{q}^{ij})$ is given by (12), and

$$\begin{aligned}[\overline{\Sigma}(t,i)\overline{\Sigma}'(t,i)]^{\imath\imath}&=\sum_{j=1}^{m_i}\nu_j^i\sum_{\jmath=1}^{d}(\sigma^{\imath\jmath}(t,\zeta_{ij}))^2,\\ \overline{Q}f(t,y,\cdot)(i)&=\sum_{j\neq i}^{l}\overline{q}^{ij}(f(t,y,j)-f(t,y,i)).\end{aligned}\tag{23}$$

Theorem 4.2: Assume (A1)–(A3). Then $(S^0(\cdot),\overline{\alpha}(\cdot))$ and $(Y^\imath(\cdot),\overline{\alpha}(\cdot))$ for $\imath=1,\dots,d$, the weak limits of $(S^{\varepsilon,0}(\cdot),\overline{\alpha}^{\varepsilon}(\cdot))$ and $(Y^{\varepsilon,\imath}(\cdot),\overline{\alpha}^{\varepsilon}(\cdot))$ are the unique solutions of the martingale problems with the operators $(\partial/\partial t)+\mathcal{L}^\imath$ for $\imath=0,\dots,d$, where $\mathcal{L}^\imath$ is defined in (22).

Proof. It is in the appendix. □

Corollary 4.3: Under the conditions of Theorem 4.2, $(S^{\varepsilon,\imath}(\cdot),\overline{\alpha}^{\varepsilon}(\cdot))$ converges weakly to $(S^\imath(\cdot),\overline{\alpha}(\cdot))$ which are solutions of

$$\begin{aligned}&\frac{dS^0(t)}{dt}=\overline{r}(t,\overline{\alpha}(t))S^0(t),\ S^0(0)=S_0^0,\\ &dS^\imath(t)=\overline{b}^\imath(t,\overline{\alpha}(t))S^\imath(t)dt+\sum_{\jmath=1}^{d}\overline{\sigma}^{\imath\jmath}(t,\overline{\alpha}(t))S^\imath(t)dw^\jmath(t),\\ &S^\imath(0)=S_0^\imath,\ \imath=1,\dots,d,\end{aligned}\tag{24}$$

respectively, where $w^\imath(\cdot)$ for $\imath=1,\dots,d$ are independent, scalar, standard Brownian motions.

Proof. Since $S^{\varepsilon,\imath}(t)=S_0^\imath\exp(Y^{\varepsilon,\imath}(t))$, and $\exp(y)$ is a continuous function, by the well-known continuous mapping theorem [2, Theorem 5.1, p.31], $S^{\varepsilon,\imath}(\cdot)$ converges to $S^\imath(\cdot)$ such that $S^\imath(t)=S_0^\imath\exp(Y^\imath(t))$, where $Y^\imath(\cdot)$ is the limit of $Y^{\varepsilon,\imath}(\cdot)$. Applying Itô's formula to (21) and noticing the relations (1) and (19), we obtain the desired result. □

Now we define the (continuous-time) limit problem of mean–variance portfolio selection. Denote by $\mathcal{F}_t$ the σ-algebra generated by $\{\overline{\alpha}(s),w(s):0\le s\le t\}$, where $w(s)=(w^1(s),\dots,w^d(s))'$. A control $u(\cdot)=(u^1(\cdot),\cdots,u^d(\cdot))$ is admissible for the limit problem if $u(\cdot)$ is $\mathcal{F}_t$-adapted and

$$\begin{aligned}dx(t)&=\{\overline{r}(t,\overline{\alpha}(t))x(t)\\ &\quad+\sum_{\imath=1}^{d}[\overline{b}^\imath(t,\overline{\alpha}(t))-\overline{r}(t,\overline{\alpha}(t))]u^\imath(t)\}dt\\ &\quad+\sum_{\jmath=1}^{d}\sum_{\imath=1}^{d}\overline{\sigma}^{\imath\jmath}(t,\overline{\alpha}(t))u^\imath(t)dw^\jmath(t)\end{aligned}\tag{25}$$

has a unique solution $x(\cdot)$ corresponding to $u(\cdot)$. Here $x(t)$ is the total wealth of the investor and $u(t)$ the allocation of wealth to the stocks, both at time t. Thus with $u^0(t)$ being the amount of fund in the bond, we must have

$$x(t)=\sum_{\imath=0}^{d}N^\imath(t)S^\imath(t)\equiv\sum_{\imath=0}^{d}u^\imath(t),\tag{26}$$

where $N^\imath(t)$ is the number of shares of the $\imath$th stock held by the investor at time t. The $(x(\cdot),u(\cdot))$ is termed an admissible wealth–portfolio pair (for the limit problem). Denote the class of admissible wealth–portfolio pairs by $\bar{\mathcal{A}}$. The limit mean–variance portfolio selection problem is formulated as

$$\begin{array}{l} \text{Minimize} \quad J(x_0, i_0, u(\cdot)) = E[x(T)-z]^2 \\ \text{subject to: } x(0)=x_0, \ \ \overline{\alpha}(0)=i_0, \quad Ex(T)=z \\ \text{and} \quad (x(\cdot),u(\cdot)) \in \overline{\mathcal{A}} \end{array} \tag{27}$$

where i_0 is such that $\ell_0 \in \mathcal{M}_{i_0}$ (recall that ℓ_0 is the initial condition of the Markov chain α_0^ε in the original discrete-time problem (6)).

For any $(x_\cdot^\varepsilon, u_\cdot^\varepsilon) \in \mathcal{A}^\varepsilon$, there are the corresponding interpolated processes $(x^\varepsilon(\cdot), u^\varepsilon(\cdot))$ determined by (10). From now on we will not distinguish between $(x_\cdot^\varepsilon, u_\cdot^\varepsilon)$ and $(x^\varepsilon(\cdot), u^\varepsilon(\cdot))$, and will occasionally write $(x^\varepsilon(\cdot), u^\varepsilon(\cdot)) \in \mathcal{A}^\varepsilon$. As a consequence of Corollary 4.3, the following result holds, which indicates that associated with the discrete-time problem (6) there is a limit continuous-time problem (27).

Corollary 4.4: Under the conditions of Theorem 4.2, any $(x^\varepsilon(\cdot), u^\varepsilon(\cdot)) \in \mathcal{A}^\varepsilon$ converges weakly to $(x(\cdot), u(\cdot))$ that belongs to $\bar{\mathcal{A}}$. Moreover, as $\varepsilon \to 0$,

$$Ex^\varepsilon(T) \to Ex(T), \quad \text{and } E[x^\varepsilon(T)-z]^2 \to E[x(T)-z]^2. \tag{28}$$

Proof. Set $\widetilde{S}^\varepsilon(\cdot) = (S^{\varepsilon,0}(\cdot), S^{\varepsilon,1}(\cdot), \ldots, S^{\varepsilon,d}(\cdot))'$. Corollary 4.3 implies that $(\widetilde{S}^\varepsilon(\cdot), \overline{\alpha}^\varepsilon(\cdot))$ converges weakly to $(\widetilde{S}(\cdot), \overline{\alpha}(\cdot)) = (S^0(t), \ldots, S^d(t), \overline{\alpha}(\cdot))$ that satisfies

$$\begin{aligned} d\widetilde{S}(t) = & \begin{pmatrix} \overline{r}(t,\overline{\alpha}(t)) & 0 \\ 0 & \operatorname{diag}(\overline{b}^1(t,\overline{\alpha}(t)), \ldots, \overline{b}^d(t,\overline{\alpha}(t))) \end{pmatrix} \widetilde{S}(t)dt \\ & + \begin{pmatrix} 0 & 0 \\ 0 & \overline{\Sigma}(t,\overline{\alpha}(t)) \operatorname{diag}(dw^1(t), \ldots, dw^d(t)) \end{pmatrix} \widetilde{S}(t). \end{aligned} \tag{29}$$

Recall that $u^\varepsilon(t) = u_k^\varepsilon$ for $t \in [\varepsilon k, \varepsilon k + \varepsilon)$. That is, $u^\varepsilon(t) = (N^1(t)S^{\varepsilon,1}(t), \ldots, N^d(t)S^{\varepsilon,d}(t))$. The weak convergence of $(S^{\varepsilon,0}(\cdot), \ldots, S^{\varepsilon,d}(\cdot))$ yields that $u^\varepsilon(\cdot)$ converges weakly to $u(\cdot)$ where $u(t) = (N^1(t)S^1(t), \ldots, N^d(t)S^d(t))$. It follows then from (5) and (26) that $x^\varepsilon(\cdot)$ converges weakly to $x(\cdot)$, which satisfies (25). This also implies $(x(\cdot), u(\cdot)) \in \bar{\mathcal{A}}$. Finally, the interpolation of x_k^ε, the weak convergence of $x^\varepsilon(\cdot)$ to $x(\cdot)$, the Skorohod representation, and the dominated convergence theorem lead to (28). □

V. Nearly Efficient Portfolio

We have established that associated with the original mean-variance control problem, there is a limit problem. In this section, we demonstrate that we can construct a nearly optimal portfolio for (6) based on the optimal portfolio for (27).

A. Efficient Portfolio of the Limit Problem

Recall that (27) is *feasible* (for fixed x_0 and i_0) if there is at least one portfolio satisfying all the constraints. It is *finite* if it is feasible and the infimum of J is finite. An optimal portfolio for a given z, if it exists, is called an *efficient portfolio*, and the corresponding $(\text{Var } x(T), z)$ is called an *efficient point*. The set of all the efficient points as z varies is termed the *efficient frontier*. In terms of the control theory terminology, an efficient portfolio is an optimal control policy (corresponding to a particular z).

In [32], a necessary and sufficient condition for the feasibility of the limit problem (27) is derived. In addition, it is proved that if it is feasible, then indeed the efficient portfolio corresponding to z exists, which can be expressed in a feedback form, namely, a function of the time t, the wealth level x and the market mode i. With the initial data given by $x(0) = x_0$ and $\overline{\alpha}(0) = i_0$, the efficient portfolio (optimal control) is given by

$$u^*(t,x,i) = -[\overline{\Sigma}(t,i)\overline{\Sigma}(t,i)']^{-1} B(t,i)'[x + (\lambda^* - z)H(t,i)], \tag{30}$$

where for $i = 1, 2, \ldots, l$,

$$\begin{aligned} &\lambda^* - z = \frac{z - P(0,i_0)H(0,i_0)x_0}{P(0,i_0)H(0,i_0)^2 + \theta - 1}, \\ &\theta = \sum_{i=1}^{l}\sum_{j=1}^{l}\int_0^T P(t,j)p_{i_0 i}(t)q_{ij}[H(t,j)-H(t,i)]^2 dt \geq 0, \\ &B(t,i) := (b_1(t,i)-r(t,i), \ldots, b_d(t,i)-r(t,i)), \\ &\dot{P}(t,i) = [\rho(t,i) - 2r(t,i)]\, P(t,i) - \sum_{j=1}^{l} q_{ij}P(t,j), \\ &P(T,i) = 1, \\ &\rho(t,i) := B(t,i)[\overline{\Sigma}(t,i)\overline{\Sigma}(t,i)']^{-1}B(t,i)', \\ &\dot{H}(t,i) = r(t,i)H(t,i) \\ &\quad -\frac{1}{P(t,i)}\sum_{j=1}^{l} q_{ij}P(t,j)[H(t,j)-H(t,i)], \ \ H(T,i) = 1. \end{aligned} \tag{31}$$

Moreover, the optimal value of Var $x(T)$, among all the wealth processes $x(\cdot)$ satisfying $Ex(T) = z$, is

$$\begin{aligned} \text{Var } x^*(T) = & \frac{P(0,i_0)H(0,i_0)^2 + \theta}{1 - \theta - P(0,i_0)H(0,i_0)^2} \\ & \cdot [z - \frac{P(0,i_0)H(0,i_0)}{P(0,i_0)H(0,i_0)^2 + \theta} x_0]^2 + \frac{P(0,i_0)\theta x_0^2}{P(0,i_0)H(0,i_0)^2 + \theta}, \end{aligned} \tag{32}$$

which gives the closed-form efficient frontier. In what follows, using an efficient portfolio of the limit problem, we construct a nearly efficient portfolio (i.e., a near-optimal control) for the original problem.

B. Nearly Efficient Portfolio

With the optimal control of the limit problem, $u^*(t,x,i)$, given by (30), we construct

$$\widetilde{u}(t,x,\alpha) = \sum_{i=1}^{l} u^*(t,x,i) I_{\{\alpha \in \mathcal{M}_i\}}. \tag{33}$$

That is, $\widetilde{u}(t,x,\alpha) = u^*(t,x,i)$ if $\alpha \in \mathcal{M}_i$, for $i = 1, 2, \cdots, l$. Let x_k^ε be the wealth trajectory of the original problem (6) under the feedback control $\widetilde{u}(\varepsilon k, x, \alpha)$. Recall that $x^\varepsilon(t)$ is the continuous-time interpolation of x_k^ε and denote the continuous-time interpolation of $\widetilde{u}(\varepsilon k, x_k^\varepsilon, \alpha_k^\varepsilon)$ by $\widetilde{u}^\varepsilon(t)$. We shall show that the use of such a control leads to near optimality of (6). Write

$$\begin{aligned} v^\varepsilon(x_0, \ell_0) &= \inf_{(x_\cdot^\varepsilon, u_\cdot^\varepsilon) \in \mathcal{A}^\varepsilon} J^\varepsilon(x_0, \ell_0, u_\cdot^\varepsilon), \quad \text{and} \\ v(x_0, i_0) &= \inf_{(x(\cdot), u(\cdot)) \in \bar{\mathcal{A}}} J(x_0, i_0, u(\cdot)). \end{aligned}$$

Theorem 5.1: Suppose that the conditions of Theorem 4.2 are satisfied. Then

$$\lim_{\varepsilon\to 0} |J^\varepsilon(x_0,\ell_0,\widetilde{u}^\varepsilon(\cdot)) - v^\varepsilon(x_0,\ell_0)| = 0. \quad (34)$$

Proof. In view of the construction of $\widetilde{u}^\varepsilon(\cdot)$ in (33), the weak convergence argument as in the proof of Theorem 4.2 yields that $\widetilde{u}^\varepsilon(\cdot)$ converges weakly to $u^*(\cdot)$, and $(x^\varepsilon(\cdot),\widetilde{u}^\varepsilon(\cdot),\overline{\alpha}^\varepsilon(\cdot))$ converges weakly to $(x(\cdot),u^*(\cdot),\overline{\alpha}(\cdot))$. Then $J^\varepsilon(x_0,\ell_0,\widetilde{u}^\varepsilon(\cdot)) \to J(x_0,i_0,u^*(\cdot)) = v(x_0,i_0)$ as $\varepsilon\to 0$. Therefore,

$$J^\varepsilon(x_0,\ell_0,\widetilde{u}^\varepsilon(\cdot)) = v(x_0,i_0) + \Delta_1(\varepsilon),$$

where $\Delta_1(\varepsilon)\to 0$ as $\varepsilon\to 0$. Select an admissible control $\widehat{u}^\varepsilon(\cdot)\in\mathcal{A}^\varepsilon$ such that

$$J^\varepsilon(x_0,\ell_0,\widehat{u}^\varepsilon(\cdot)) \le v^\varepsilon(x_0,\ell_0)+\varepsilon.$$

Define

$$\overline{u}(t,x,\alpha) = \sum_{i=1}^{l} \widehat{u}^\varepsilon(t,x,i) I_{\{\alpha\in\mathcal{M}_i\}},$$

set $\overline{x}_k^\varepsilon$ as in (5) but with α_k^ε replaced by $\overline{\alpha}_k^\varepsilon$, and let $\overline{x}^\varepsilon(\cdot)$ and $\overline{u}^\varepsilon(\cdot)$ be the piecewise constant interpolations of $\overline{x}_k^\varepsilon$ and $\overline{u}(\varepsilon k,\overline{x}_k^\varepsilon,\overline{\alpha}_k^\varepsilon)$, respectively. Then $J^\varepsilon(x_0,\ell_0,\overline{u}^\varepsilon(\cdot)) = E[\overline{x}^\varepsilon(T)-z]^2$. Similar to the argument leading to the equation after (49) in [14], using the mean squares estimate (15) on the occupation measure, the wealth equation (5), the definition of $\overline{x}_k^\varepsilon$, and the Gronwall's inequality, we can show that $E|x^\varepsilon(t)-\overline{x}^\varepsilon(t)|^2\to 0$ as $\varepsilon\to 0$ for $t\in[0,T]$. This implies that

$$\begin{aligned} J^\varepsilon(x_0,\ell_0,\overline{u}^\varepsilon(\cdot)) &\le J^\varepsilon(x_0,\ell_0,\widehat{u}^\varepsilon(\cdot)) + \Delta_2(\varepsilon) \\ &\le v^\varepsilon(x_0,\ell_0)+\varepsilon+\Delta_2(\varepsilon), \end{aligned} \quad (35)$$

where $\Delta_2(\varepsilon)\to 0$ as $\varepsilon\to 0$. The tightness of $(x^\varepsilon(\cdot),\overline{u}^\varepsilon(\cdot),\overline{\alpha}^\varepsilon(\cdot))$ implies that we can extract a convergent subsequence, still denoted by $(x^\varepsilon(\cdot),\overline{u}^\varepsilon(\cdot),\overline{\alpha}^\varepsilon(\cdot))$ for simplicity, such that $J^\varepsilon(x_0,\ell_0,\overline{u}^\varepsilon(\cdot))\to J(x_0,i_0,\overline{u}(\cdot))$. It follows that

$$v(x_0,i_0) \le J(x_0,i_0,\overline{u}(\cdot)) = J^\varepsilon(x_0,\ell_0,\overline{u}^\varepsilon(\cdot)) + \Delta_3(\varepsilon), \quad (36)$$

where $\Delta_3(\varepsilon)\to 0$ as $\varepsilon\to 0$. Combining (35) and (36),

$$\begin{aligned} v^\varepsilon(x_0,\ell_0) &\le J^\varepsilon(x_0,\ell_0,\widetilde{u}^\varepsilon(\cdot)) \\ &= v(x_0,i_0)+\Delta_1(\varepsilon) \\ &\le J(x_0,i_0,\overline{u}(\cdot))+\Delta_1(\varepsilon) \\ &= J^\varepsilon(x_0,\ell_0,\overline{u}^\varepsilon(\cdot))+\Delta_1(\varepsilon)+\Delta_3(\varepsilon) \\ &\le v^\varepsilon(x_0,\ell_0)+\varepsilon+\Delta_1(\varepsilon)+\Delta_2(\varepsilon)+\Delta_3(\varepsilon). \end{aligned} \quad (37)$$

Subtracting $v^\varepsilon(x_0,\ell_0)$ from (37) and taking the limit as $\varepsilon\to 0$, we arrive at (34). □

VI. EXTENSIONS

We have demonstrated the natural connections of the discrete-time model and its continuous-time counterpart for the Markowitz's mean-variance portfolio selection problems under the premise that the modulating Markov chain is time homogeneous. In this section, we extend the results to time inhomogeneous Markov chains, which entail the time-dependent transition matrices. To proceed, we first recall the definition of weak irreducibility.

Definition 6.1: Suppose that α_k is a discrete-time Markov chain with a finite state space $\mathcal{M}=\{1,\dots,m\}$ and (time-dependent) transition probability matrices P_k. The Markov chain (or the transition matrix) is said to be weakly irreducible if for each k, the system of equations

$$\nu_k P_k = \nu_k, \quad \nu_k \mathbb{1}_m = 1 \quad (38)$$

has a unique solution $\nu_k = (\nu_{k,1},\dots,\nu_{k,m}) \in \mathbb{R}^{1\times m}$ and $\nu_{k,i}\ge 0$ for each $i=1,\dots,m$. The unique nonnegative solution is termed a quasi-stationary distribution.

Remark 6.2: The definition above is an extension of the usual notion of irreducibility. It allows the transition matrices and the quasi-stationary measures to be time dependent. In addition, some of the components $\nu_{k,i}$ could be 0.

Using Definition 6.1, we modify the argument used in the previous sections, and extend the results to include time-dependent transition probabilities. Since the proofs of the assertions are similar to the previous case, we will only present the conditions needed and the final results. The verbatim proofs will be omitted. In lieu of (A1), we use:

(A1') Suppose that α_k^ε is a Markov chain with state space $\mathcal{M}$ and time-dependent transition matrices

$$P_k^\varepsilon = P(\varepsilon k)+\varepsilon Q(\varepsilon k) \quad (39)$$

such that $P(\varepsilon k)$ has the form of the decomposition as in (8) with each P^i replaced by $P^i(\varepsilon k)$. For each $i=1,\dots,l$, $0\le k\le T/\varepsilon$, $P^i(\varepsilon k)$ is weakly irreducible and aperiodic. As functions of $t\in[0,T]$, $P(\cdot)$ is continuously differentiable with Lipschitz continuous derivative, and $Q(\cdot)$ is Lipschitz continuous.

Theorem 6.3: Under the conditions of Theorem 4.2 with (A1) replaced by (A1'), the conclusions of Theorem 4.2 and Corollary 4.4 continue to hold with ν^i and $\overline{Q}$ replaced by $\nu^i(t)$ and $\overline{Q}(t)$, respectively.

Since in a finite-state Markov chain, in addition to recurrent states, transient states may also be included. Our next result concerns about the inclusion of transient states in addition to the l classes of weakly irreducible classes. In this case we need to replace (A1') by (A1").

(A1") Suppose that the Markov chain α_k^ε has a finite state space $\mathcal{M}$ and transition matrix (39) with

$$P(\varepsilon k) = \begin{pmatrix} P^1(\varepsilon k) & & & & \\ & P^2(\varepsilon k) & & & \\ & & \ddots & & \\ & & & P^l(\varepsilon k) & \\ P^{*,1}(\varepsilon k) & P^{*,2}(\varepsilon k) & \cdots & P^{*,l}(\varepsilon k) & P^*(\varepsilon k) \end{pmatrix}. \quad (40)$$

For each $0\le k\le T/\varepsilon$, and each $i=1,\dots,l$, the transition probability matrices $P^i(\varepsilon k)$ are weakly irreducible and aperiodic. For each $0\le k\le T/\varepsilon$, $P^*(\varepsilon k)$ is a matrix having all of its eigenvalues inside the unit circle. As functions of $t\in[0,T]$, $P(\cdot)$ is continuously differentiable with Lipschitz continuous derivative, and $Q(\cdot)$ is Lipschitz continuous. In addition, with t denoting εk, there exists an $m_*\times m_*$ nonsingular matrix $B(t)$ and constant matrices P^* and $P^{*,i}$ satisfying $P^*(t) - I = B(t)(P^* - I)$, and $P^{*,i}(t) = B(t)P^{*,i}$, for $i = 1, \ldots, l$.

Now the state space of the Markov chain can be written as: $\mathcal{M} = \mathcal{M}_1 \cup \mathcal{M}_2 \cup \cdots \cup \mathcal{M}_l \cup \mathcal{M}_*$, where $\mathcal{M}_i = \{\zeta_{i1}, \ldots, \zeta_{im_i}\}$ for $i = 1, \ldots, l$, and $\mathcal{M}_* = \{\zeta_{*1}, \ldots, \zeta_{*m_*}\}$. Define

$$a_i(\varepsilon k) = -(P^*(\varepsilon k) - I)^{-1} P^{*,i}(\varepsilon k) \mathbb{1}_{m_i}, \text{ for } i = 1, \ldots, l.$$

Then it is easily seen that $a_i(\varepsilon k) = a_i$ is independent of k. Denote the jth component of a_i by $a_{i,j}$. Partition the matrix $Q(t)$ as

$$Q(t) = \begin{pmatrix} Q^{11}(t) & Q^{12}(t) \\ Q^{21}(t) & Q^{22}(t) \end{pmatrix}$$

where $Q^{11}(t) \in \mathbb{R}^{(m-m_*)\times(m-m_*)}$, $Q^{12}(t) \in \mathbb{R}^{(m-m_*)\times m_*}$, $Q^{21}(t) \in \mathbb{R}^{m_* \times (m-m_*)}$, and $Q^{22}(t) \in \mathbb{R}^{m_* \times m_*}$, and define

$$\overline{Q}_*(t) = \text{diag}(\nu^1(t), \ldots, \nu^l(t))(Q^{11}(t)\widetilde{\mathbb{1}} + Q^{12}(t)(a_1, \ldots, a_l). \tag{41}$$

In carrying out the aggregation, we only lump all the states in each weakly irreducible class into a single state. This leads to the definition of the aggregated process. Let U be a random variable uniformly distributed in $[0, 1]$ and let

$$\begin{aligned} U_j \ &= I_{\{0 \le U \le a_{1,j}\}} + 2I_{\{a_{1,j} < U \le a_{1,j} + a_{2,j}\}} + \cdots \\ &\quad + lI_{\{a_{1,j} + \cdots + a_{l-1,j} < U \le 1\}}. \end{aligned}$$

Then define $\overline{\alpha}^\varepsilon_k = \begin{cases} i & \text{if } \alpha^\varepsilon_k \in \mathcal{M}_i, \\ U_j, & \text{if } \alpha^\varepsilon_k = \zeta_{*j}. \end{cases}$ Using essentially the same idea but more complex and detailed estimates, we can derive the following results.

Theorem 6.4: Under the conditions of Theorem 4.2 with (A1) replaced by (A1”), the conclusions of Theorem 4.2 and Corollary 4.4 continue to hold with ν^i and $\overline{Q}$ replaced by the time-varying $\nu^i(t)$ and $\overline{Q}_*(t)$ given by (41), respectively.

VII. Concluding Remarks

This paper has been devoted to a class of dynamic Markowitz’s mean-variance portfolio selection problems. Taking into consideration of market trend and other factors, a discrete-time model that is modulated by a Markov chain was introduced. Aiming at complexity reduction, we use nearly completely decomposable transition matrices and weak convergence methods to derive the limit mean-variance portfolio selection problem. Based on the limit, we can design optimal (efficient) portfolios and derive efficient frontier [32] (see also the framework of LQ control with indefinite control weights [4]). Then using the efficient portfolios of the limit problem, we constructed portfolios for the original discrete-time model and show that such portfolios are nearly efficient.

In Section 6, we assumed the smoothness of the transition functions in (A1’) and (A1”). Such a condition can be relaxed if we work with convergence of the probability vector p^ε_k and the transition probability matrices under the weak topology of $L_2[0, T]$. The associated weak convergence and the limit systems can still be obtained. As far as the limit mean-variance problem is concerned, for the cases discussed in Section 6, due to the time-dependent generator (the non-homogeneity), the corresponding limit problem is more difficult to handle. However, we can still obtain near optimality if we use a “δ-optimal” portfolio policy for the limit problem. Based on such a δ-optimal policy, we can construct portfolios that are nearly optimal for the original problem. The statement of (34) is changed to $\limsup_{\varepsilon\to 0} |J^\varepsilon(x, \ell_0, u^\varepsilon(\cdot)) - v^\varepsilon(x, \ell_0)| \le \delta$.

This paper has been devoted to a discrete-time model. In fact, one may use the weak convergence method to treat a continuous-time mean-variance portfolio problem:

$$\begin{aligned} \frac{dS^{\varepsilon,0}(t)}{dt} &= r(t, \alpha^\varepsilon(t))S^{\varepsilon,0}(t) \\ dS^{\varepsilon,\imath}(t) &= S^{\varepsilon,\imath}(t)[b(t, \overline{\alpha}^\varepsilon(t))dt + \sum_{\jmath=1}^{d} \sigma^{\imath\jmath}(t, \alpha^\varepsilon(t))dw^\jmath(t)], \\ &\quad \imath = 1, \ldots, d, \end{aligned}$$

where $\alpha^\varepsilon(\cdot)$ is a continuous-time singularly perturbed Markov chain generated by $Q^\varepsilon(t) = \widetilde{Q}(t)/\varepsilon + \widehat{Q}(t)$. The use of the singularly perturbed Markov chain again comes from the motivation of reduction of complexity. As shown in [24, §2.2, p.840], such a continuous-time Markov chain has an associated discrete-time chain whose transition matrix has the form (7). Thus the problem treated in this paper can be thought of as a discretization of the continuous-time problem given above.

It should be noted that in our model the wealth of an agent is allowed to become and remain negative. Prohibition of bankruptcy renders the problem one with state constraint, which remains an interesting yet challenging open problem.

Finally, we remark that although the formulation and motivation in this paper stem from mean-variance portfolio selection problems, the techniques used and the methods of solutions are not restricted to financial engineering applications. They can also be employed in other hybrid control problems modulated by a singularly perturbed Markov chain.

Appendix

Proof Theorem 4.1**.** By virtue of Lemma 3.2-(iii), $\{\overline{\alpha}^\varepsilon(\cdot)\}$ is tight. In view of the well-know Crámer-Wold device [2, p. 49], to prove the tightness of $\{Y^{\varepsilon,\imath}(\cdot), \overline{\alpha}^\varepsilon(\cdot)\}$, it suffices to establish the tightness of $\{Y^{\varepsilon,\imath}(\cdot)\}$.

Use E^ε_t to denote the conditional expectation with respect to $\mathcal{F}^\varepsilon_t$, the σ-algebra generated by $\{\alpha^\varepsilon_{k_1}, \xi^\imath_{k_1} : k_1 < t/\varepsilon,\ \imath = 0, \ldots, d\}$. For any $\eta > 0$, $t \ge 0$ and $0 \le s \le \eta$, and for each $\imath = 1, \ldots, d$,

$$\begin{aligned} &E^\varepsilon_t[Y^{\varepsilon,\imath}(t+s) - Y^{\varepsilon,\imath}(t)]^2 \\ &\le KE^\varepsilon_t \left| \varepsilon \sum_{k=t/\varepsilon}^{(t+s)/\varepsilon - 1} c^\imath(\varepsilon k, \alpha^\varepsilon_k) \right|^2 \\ &\quad + KE^\varepsilon_t \left| \sqrt{\varepsilon} \sum_{k=t/\varepsilon}^{(t+s)/\varepsilon - 1} \sum_{\jmath=1}^{d} \sigma^{\imath\jmath}(\varepsilon k, \alpha^\varepsilon_k)\xi^\jmath_k \right|^2 \\ &\le K\varepsilon^2 E^\varepsilon_t \sum_{k=t/\varepsilon}^{(t+s)/\varepsilon - 1} \sum_{k_1=t/\varepsilon}^{(t+s)/\varepsilon - 1} c^\imath(\varepsilon k, \alpha^\varepsilon_k)c^\imath(\varepsilon k_1, \alpha^\varepsilon_{k_1}) \\ &\quad + K\varepsilon E^\varepsilon_t \sum_{k=t/\varepsilon}^{(t+s)/\varepsilon - 1} \left(\sum_{\jmath=1}^{d} \sigma^{\imath\jmath}(\varepsilon k, \alpha^\varepsilon_k)\xi^\jmath_k \right)^2 . \end{aligned} \tag{42}$$

In the above and hereafter, K represents a generic positive real number; its values may be different for different appearances. Note that t/ε, $(t+s)/\varepsilon$ are understood to be the integer parts in the summation limits of (42). In the above, we have also used the independence of $\{\xi_k\}$ with $\{\alpha_k^\varepsilon\}$, and $E_t^\varepsilon[\xi_k^\imath \xi_{k_1}^\imath]=0$ for $k_1 \neq k$ and $t/\varepsilon \le k_1, k \le (t+s)/\varepsilon$. The boundedness of $c^\imath(\cdot)$ then implies that

$$\varepsilon^2 E_t^\varepsilon \left| \sum_{k=t/\varepsilon}^{(t+s)/\varepsilon-1} \sum_{k_1=t/\varepsilon}^{(t+s)/\varepsilon-1} c^\imath(\varepsilon k, \alpha_k^\varepsilon) c^\imath(\varepsilon k_1, \alpha_{k_1}^\varepsilon) \right| \le K\varepsilon^2 \left(\frac{t+s}{\varepsilon} - \frac{t}{\varepsilon} \right)^2 \le K s^2.$$

Similarly, the boundedness of $\sigma^{\imath\jmath}(\cdot)$ implies that

$$\varepsilon E_t^\varepsilon \sum_{k=t/\varepsilon}^{(t+s)/\varepsilon-1} \left(\sum_{\jmath=1}^{d} \sigma^{\imath\jmath}(\varepsilon k, \alpha_k^\varepsilon) \xi_k^\jmath \right)^2 \le K s.$$

Consequently,

$$\lim_{\eta\to 0} \limsup_{\varepsilon\to 0} E[Y^{\varepsilon,\imath}(t+s) - Y^{\varepsilon,\imath}(t)]^2 = 0.$$

The desired result then follows from the tightness criterion [12, Theorem 3, p.47]. Likewise, it can be verified that $\lim_{\eta\to 0}\limsup_{\varepsilon\to 0} E[S^{\varepsilon,0}(t+s) - S^{\varepsilon,0}(t)]^2 = 0$. Thus $\{S^{\varepsilon,0}(\cdot)\}$ is also tight. □

Proof Theorem 4.2**.** We focus on the stock price processes $(Y^{\varepsilon,\imath}(\cdot), \overline{\alpha}^\varepsilon(\cdot))$ for $\imath = 1,\ldots,d$. To proceed, let us fix $\imath \in \{1,\ldots,d\}$. The uniqueness of the solution of the martingale problem can be verified as in [23, Lemma 7.18]. To characterize the limit, it suffices to show that for each $i \in \overline{\mathcal{M}}$ and $f(\cdot,\cdot,i) \in C_0^{1,2}$ (where $C_0^{1,2}$ is the collection of functions that have compact support and that are continuously differentiable with respect to the first variable, and twice continuously differentiable with respect to the second variable),

$$\begin{aligned} & f(t, Y^\imath(t), \overline{\alpha}(t)) - f(0, Y^\imath(0), \overline{\alpha}(0)) \\ & - \int_0^t \left(\frac{\partial}{\partial \tau} + \mathcal{L}^\imath \right) f(\tau, Y^\imath(\tau), \overline{\alpha}(\tau)) d\tau \text{ is a martingale.} \end{aligned} \tag{43}$$

To obtain (43), it suffices to show that for any positive integer κ, bounded and continuous function $h_{j_1}(\cdot)$ with $j_1 \le \kappa$, and any $t, s, t_{j_1} \ge 0$ satisfying $t_{j_1} \le t < t+s \le T$,

$$\begin{aligned} & E \prod_{j_1=1}^{\kappa} h_{j_1}(Y^\imath(t_{j_1}), \overline{\alpha}(t_{j_1})) \Big(f(t+s, Y^\imath(t+s), \overline{\alpha}(t+s)) \\ & - f(t, Y^\imath(t), \overline{\alpha}(t)) - \int_t^{t+s} \big(\frac{\partial}{\partial\tau} + \mathcal{L}^\imath\big) \\ & \qquad f(\tau, Y^\imath(\tau), \overline{\alpha}(\tau)) d\tau \Big) = 0. \end{aligned} \tag{44}$$

To verify (44), we begin with the pre-limit processes (the processes indexed by ε). Define

$$\overline{f}(t,y,\alpha) = \sum_{i=1}^{l} f(t,y,i) I_{\{\alpha\in\mathcal{M}_i\}} \quad \text{for each} \quad \alpha \in \mathcal{M}. \tag{45}$$

Then $\overline{f}(\cdot,\cdot,\alpha) \in C_0^{1,2}$. The weak convergence of $(Y^{\varepsilon,\imath}(\cdot), \overline{\alpha}^\varepsilon(\cdot))$ to $(Y^\imath(\cdot), \overline{\alpha}(\cdot))$, the Skorohod representation, the continuity of $f(\cdot)$, and the dominated convergence theorem then yield that as $\varepsilon \to 0$,

$$\begin{aligned} & E \prod_{j_1=1}^{\kappa} h_{j_1}(Y^\imath(t_{j_1}), \overline{\alpha}(t_{j_1})) \\ & \cdot (f(t+s, Y^{\varepsilon,\imath}(t+s), \overline{\alpha}^\varepsilon(t+s)) - f(t, Y^{\varepsilon,\imath}(t), \overline{\alpha}^\varepsilon(t))) \\ & \to E \prod_{j_1=1}^{\kappa} h_{j_1}(Y^\imath(t_{j_1}), \overline{\alpha}(t_{j_1})) \\ & \quad \cdot (f(t+s, Y^\imath(t+s), \overline{\alpha}(t+s)) - f(t, Y^\imath(t), \overline{\alpha}(t))) . \end{aligned} \tag{46}$$

Pick out $\{n_\varepsilon\}$, a sequence of positive integers such that $n_\varepsilon \to \infty$ as $\varepsilon \to 0$ and $\varepsilon n_\varepsilon = \delta_\varepsilon \to 0$ as $\varepsilon \to 0$. Then

$$\begin{aligned} & f(t+s, Y^{\varepsilon,\imath}(t+s), \overline{\alpha}^\varepsilon(t+s)) - f(t, Y^{\varepsilon,\imath}(t), \overline{\alpha}^\varepsilon(t)) \\ & = \int_t^{t+s} \frac{\partial}{\partial\tau} f(\tau, Y^{\varepsilon,\imath}(\tau), \overline{\alpha}^\varepsilon(\tau)) d\tau \\ & \quad + \sum_{l\delta_\varepsilon=t}^{t+s} [\overline{f}(l\delta_\varepsilon, Y^{\varepsilon,\imath}_{ln_\varepsilon}, \alpha^\varepsilon_{ln_\varepsilon+n_\varepsilon}) - \overline{f}(l\delta_\varepsilon, Y^{\varepsilon,\imath}_{ln_\varepsilon}, \alpha^\varepsilon_{ln_\varepsilon})] \\ & \quad + \sum_{l\delta_\varepsilon=t}^{t+s} [\overline{f}(l\delta_\varepsilon, Y^{\varepsilon,\imath}_{ln_\varepsilon+n_\varepsilon}, \alpha^\varepsilon_{ln_\varepsilon+n_\varepsilon}) - \overline{f}(l\delta_\varepsilon, Y^{\varepsilon,\imath}_{ln_\varepsilon}, \alpha^\varepsilon_{ln_\varepsilon+n_\varepsilon})] \\ & \quad + o(1), \end{aligned} \tag{47}$$

where $o(1) \to 0$ in probability as $\varepsilon \to 0$ uniformly in $t \in [0,T]$. Note that we have used $\overline{f}(t, Y_k^{\varepsilon,\imath}, \alpha_k^\varepsilon) = f(t, Y_k^{\varepsilon,\imath}, \overline{\alpha}_k^\varepsilon)$ in the above.

By virtue of the weak convergence, the Skorohod representation, and the continuity and the boundedness of $(\partial/\partial\tau) f(\cdot,\cdot,\alpha)$, as $\varepsilon \to 0$,

$$\begin{aligned} & E \prod_{j_1=1}^{\kappa} h_{j_1}(Y^{\varepsilon,\imath}(t_{j_1}), \overline{\alpha}^\varepsilon(t_{j_1})) \\ & \cdot \big(\int_t^{t+s} \frac{\partial}{\partial\tau} f(\tau, Y^{\varepsilon,\imath}(\tau), \overline{\alpha}^\varepsilon(\tau)) d\tau \big) \\ & \to E \prod_{j_1=1}^{\kappa} h_{j_1}(Y^\imath(t_{j_1}), \overline{\alpha}(t_{j_1})) \\ & \cdot \big(\int_t^{t+s} \frac{\partial}{\partial\tau} f(\tau, Y^\imath(\tau), \overline{\alpha}(\tau)) d\tau \big). \end{aligned} \tag{48}$$

As for the terms on the next to the last line of (47), define $\widehat{I}(\alpha) = (I_{\{\alpha=\zeta_{ij}\}}, 1 \le i \le l, 1 \le j \le m_i)$, $\overline{F}(t,y) = \begin{pmatrix} f(t,y,1)\mathbb{1}_{m_1} \\ \vdots \\ f(t,y,l)\mathbb{1}_{m_l} \end{pmatrix}$. Since $Y^{\varepsilon,\imath}_{ln_\varepsilon}$ is $\mathcal{F}^\varepsilon_{k+1}$-measurable for $ln_\varepsilon \le k < ln_\varepsilon + n_\varepsilon$, and $P^\varepsilon = P + \varepsilon Q$ given by (7),

$$\begin{aligned}
&E_t^\varepsilon \sum_{l\delta_\varepsilon=t}^{t+s} [\overline{f}(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}, \alpha_{ln_\varepsilon+n_\varepsilon}^\varepsilon) - \overline{f}(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}, \alpha_{ln_\varepsilon}^\varepsilon)] \\
&= E_t^\varepsilon \sum_{l\delta_\varepsilon=t}^{t+s} \sum_{k=ln_\varepsilon}^{ln_\varepsilon+n_\varepsilon-1} E_{k+1}[\overline{f}(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}, \alpha_{k+1}^\varepsilon) \\
&\quad -\overline{f}(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}, \alpha_k^\varepsilon)] \\
&= E_t^\varepsilon \sum_{l\delta_\varepsilon=t}^{t+s} \sum_{k=ln_\varepsilon}^{ln_\varepsilon+n_\varepsilon-1} \widehat{I}(\alpha_k^\varepsilon)(P^\varepsilon - I)\overline{F}(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}) \\
&= \varepsilon E_t^\varepsilon \sum_{l\delta_\varepsilon=t}^{t+s} \sum_{k=ln_\varepsilon}^{ln_\varepsilon+n_\varepsilon-1} \widehat{I}(\alpha_k^\varepsilon) Q\overline{F}(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}) \\
&= \varepsilon E_t^\varepsilon \sum_{l\delta_\varepsilon=t}^{t+s} \sum_{k=ln_\varepsilon}^{ln_\varepsilon+n_\varepsilon-1} Q\overline{f}(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}, \cdot)(\alpha_k^\varepsilon).
\end{aligned} \tag{49}$$

In (49), we have used the orthogonality $(P - I)\overline{F}(t, x) = 0$. Furthermore,

$$\begin{aligned}
&\varepsilon \sum_{l\delta_\varepsilon=t}^{t+s} \sum_{k=ln_\varepsilon}^{ln_\varepsilon+n_\varepsilon-1} Q\overline{f}(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}, \cdot)(\alpha_k^\varepsilon) \\
&= \sum_{l\delta_\varepsilon=t}^{t+s} \sum_{i=1}^{l} \sum_{j=1}^{m_i} Q\overline{f}(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}, \cdot)(\zeta_{ij})\delta_\varepsilon \\
&\qquad \cdot \frac{1}{n_\varepsilon} \sum_{k=ln_\varepsilon}^{ln_\varepsilon+n_\varepsilon-1} I_{\{\alpha_k^\varepsilon=\zeta_{ij}\}} \\
&= \sum_{l\delta_\varepsilon=t}^{t+s} \sum_{i=1}^{l} \sum_{j=1}^{m_i} Q\overline{f}(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}, \cdot)(\zeta_{ij})\delta_\varepsilon \\
&\qquad \cdot \frac{1}{n_\varepsilon} \sum_{k=ln_\varepsilon}^{ln_\varepsilon+n_\varepsilon-1} \nu_j^i I_{\{\alpha_k^\varepsilon\in\mathcal{M}_i\}} \\
&+ \sum_{l\delta_\varepsilon=t}^{t+s} \sum_{i=1}^{l} \sum_{j=1}^{m_i} Q\overline{f}(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}, \cdot)(\zeta_{ij})\delta_\varepsilon \\
&\cdot \frac{1}{n_\varepsilon} \sum_{k=ln_\varepsilon}^{ln_\varepsilon+n_\varepsilon-1} [I_{\{\alpha_k^\varepsilon=\zeta_{ij}\}} - \nu_j^i I_{\{\overline{\alpha}_k=i\}}].
\end{aligned}$$

Lemma 3.2 implies that the last line above goes to 0 in probability uniformly in $t \in [0, T]$. In view of (12), as $\varepsilon \to 0$,

$$\begin{aligned}
E\Big| \sum_{l\delta_\varepsilon=t}^{t+s} \delta_\varepsilon \frac{1}{n_\varepsilon} \sum_{k=ln_\varepsilon}^{ln_\varepsilon+n_\varepsilon-1} &[Q\overline{f}(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}, \cdot)(\alpha_k^\varepsilon) \\
&-\overline{Q}f(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}, \cdot)(\overline{\alpha}_k^\varepsilon)]\Big| \to 0
\end{aligned}$$

uniformly in t. Putting the above estimates together, as $\varepsilon \to 0$,

$$\begin{aligned}
&E \prod_{j_1=1}^{\kappa} h_{j_1}(Y^{\varepsilon,\imath}(t_{j_1}), \overline{\alpha}^\varepsilon(t_{j_1})) \\
&\cdot\Big(\sum_{l\delta_\varepsilon=t}^{t+s} \delta_\varepsilon \frac{1}{n_\varepsilon} \sum_{k=ln_\varepsilon}^{ln_\varepsilon+n_\varepsilon-1} Q\overline{f}(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}, \cdot)(\alpha_k^\varepsilon)\Big) \\
&\to E \prod_{j_1=1}^{\kappa} h_{j_1}(Y^\imath(t_{j_1}), \overline{\alpha}(t_{j_1})) \\
&\quad \cdot\Big(\int_t^{t+s} \overline{Q}f(\tau, Y^\imath(\tau), \cdot)(\overline{\alpha}(\tau))d\tau\Big).
\end{aligned} \tag{50}$$

Using a truncated Taylor expansion and denoting $\overline{f}_y = (\partial/\partial y)\overline{f}$ and $\overline{f}_{yy} = (\partial^2/\partial y^2)\overline{f}$, we have

$$\begin{aligned}
&\sum_{l\delta_\varepsilon=t}^{t+s} [\overline{f}(l\delta_\varepsilon, Y_{ln_\varepsilon+n_\varepsilon}^{\varepsilon,\imath}, \alpha_{ln_\varepsilon+n_\varepsilon}^\varepsilon) - \overline{f}(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}, \alpha_{ln_\varepsilon+n_\varepsilon}^\varepsilon)] \\
&= \sum_{l\delta_\varepsilon=t}^{t+s} \overline{f}_y(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}, \alpha_{ln_\varepsilon+n_\varepsilon}^\varepsilon)[Y_{ln_\varepsilon+n_\varepsilon}^{\varepsilon,\imath} - Y_{ln_\varepsilon}^{\varepsilon,\imath}] \\
&+\frac{1}{2} \sum_{l\delta_\varepsilon=t}^{t+s} \overline{f}_{yy}(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}, \alpha_{ln_\varepsilon+n_\varepsilon}^\varepsilon)[Y_{ln_\varepsilon+n_\varepsilon}^{\varepsilon,\imath} - Y_{ln_\varepsilon}^{\varepsilon,\imath}]^2 \\
&+\frac{1}{2} \sum_{l\delta_\varepsilon=t}^{t+s} [\overline{f}_{yy}(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath,+}, \alpha_{ln_\varepsilon+n_\varepsilon}^\varepsilon) \\
&\qquad -\overline{f}_{yy}(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}, \alpha_{ln_\varepsilon+n_\varepsilon}^\varepsilon)][Y_{ln_\varepsilon+n_\varepsilon}^{\varepsilon,\imath} - Y_{ln_\varepsilon}^{\varepsilon,\imath}]^2,
\end{aligned} \tag{51}$$

where $Y_{ln_\varepsilon}^{\varepsilon,\imath,+}$ is on the line segment joining $Y_{ln_\varepsilon}^{\varepsilon,\imath}$ and $Y_{ln_\varepsilon+n_\varepsilon}^{\varepsilon,\imath}$. Substituting (16) in (51), we can write

$$\begin{aligned}
&\sum_{l\delta_\varepsilon=t}^{t+s} \overline{f}_y(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}, \alpha_{ln_\varepsilon+n_\varepsilon}^\varepsilon)[Y_{ln_\varepsilon+n_\varepsilon}^{\varepsilon,\imath} - Y_{ln_\varepsilon}^{\varepsilon,\imath}] \\
&= \sum_{l\delta_\varepsilon=t}^{t+s} \overline{f}_y(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}, \alpha_{ln_\varepsilon+n_\varepsilon}^\varepsilon) \sum_{k=ln_\varepsilon}^{ln_\varepsilon+n_\varepsilon-1} \varepsilon c^\imath(\varepsilon k, \alpha_k^\varepsilon) \\
&+ \sum_{l\delta_\varepsilon=t}^{t+s} \overline{f}_y(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}, \alpha_{ln_\varepsilon+n_\varepsilon}^\varepsilon) \\
&\cdot \sum_{k=ln_\varepsilon}^{ln_\varepsilon+n_\varepsilon-1} \sqrt{\varepsilon} \sum_{\jmath=1}^{d} \sigma^{\imath\jmath}(\varepsilon k, \alpha_k^\varepsilon)\xi_k^\jmath.
\end{aligned} \tag{52}$$

To proceed, we first replace $\alpha_{ln_\varepsilon+n_\varepsilon}^\varepsilon$ in the argument of $\overline{f}_y(\cdot)$ above by $\alpha_{ln_\varepsilon}^\varepsilon$. Letting $l\delta_\varepsilon \to \tau$ as $\varepsilon \to 0$, and using Lemma 3.2-(iii), for all $ln_\varepsilon \le k \le ln_\varepsilon + n_\varepsilon - 1$, $\varepsilon k \to \tau$,

$$\begin{aligned}
&E \prod_{j_1=1}^{\kappa} h_{j_1}(Y^{\varepsilon,\imath}(t_{j_1}), \overline{\alpha}^\varepsilon(t_{j_1})) \sum_{l\delta_\varepsilon=t}^{t+s} \overline{f}_y(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}, \alpha_{ln_\varepsilon}^\varepsilon) \\
&\cdot\delta_\varepsilon \frac{1}{n_\varepsilon} \sum_{k=ln_\varepsilon}^{ln_\varepsilon+n_\varepsilon-1} c^\imath(\varepsilon k, \alpha_k^\varepsilon) \\
&= E \prod_{j_1=1}^{\kappa} h_{j_1}(Y^{\varepsilon,\imath}(t_{j_1}), \overline{\alpha}^\varepsilon(t_{j_1})) \sum_{l\delta_\varepsilon=t}^{t+s} \overline{f}_y(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}, \alpha_{ln_\varepsilon}^\varepsilon) \\
&\cdot\frac{\delta_\varepsilon}{n_\varepsilon} \sum_{k=ln_\varepsilon}^{ln_\varepsilon+n_\varepsilon-1} \sum_{i=1}^{l} \sum_{j=1}^{m_i} c^\imath(\varepsilon k, \zeta_{ij})\nu_j^i I_{\{\alpha_k^\varepsilon\in\mathcal{M}_i\}} \\
&+E \prod_{j_1=1}^{\kappa} h_{j_1}(Y^{\varepsilon,\imath}(t_{j_1}), \overline{\alpha}^\varepsilon(t_{j_1})) \sum_{l\delta_\varepsilon=t}^{t+s} \overline{f}_y(l\delta_\varepsilon, Y_{ln_\varepsilon}^{\varepsilon,\imath}, \alpha_{ln_\varepsilon}^\varepsilon) \\
&\cdot\delta_\varepsilon \frac{1}{n_\varepsilon} \sum_{k=ln_\varepsilon}^{ln_\varepsilon+n_\varepsilon-1} \sum_{i=1}^{l} \sum_{j=1}^{m_i} c^\imath(\varepsilon k, \zeta_{ij}) \\
&\cdot[I_{\{\alpha_k^\varepsilon=\zeta_{ij}\}} - \nu_j^i I_{\{\alpha_k^\varepsilon\in\mathcal{M}_i\}}] \\
&\to E \prod_{j_1=1}^{\kappa} h_{j_1}(Y^\imath(t_{j_1}), \overline{\alpha}(t_{j_1})) \\
&\cdot \int_t^{t+s} f_y(\tau, Y^\imath(\tau), \overline{\alpha}(\tau))\overline{c}^\imath(\tau, \overline{\alpha}(\tau))d\tau.
\end{aligned} \tag{53}$$

By virtue of the independence of $\{\xi_k^\imath\}$, inserting E_t^ε and then

E_k, and using $E_k\xi_k^{\jmath}=0$, we have

$$E\prod_{j_1=1}^{\kappa} h_{j_1}(Y^{\varepsilon,\imath}(t_{j_1}),\overline{\alpha}^{\varepsilon}(t_{j_1}))\Big[\sum_{l\delta_\varepsilon=t}^{t+s}\overline{f}_y(l\delta_\varepsilon,Y^{\varepsilon,\imath}_{ln_\varepsilon},\alpha^{\varepsilon}_{ln_\varepsilon}) \cdot\sum_{k=ln_\varepsilon}^{ln_\varepsilon+n_\varepsilon-1}\sqrt{\varepsilon}\sigma^{\imath\jmath}(\varepsilon k,\alpha_k^{\varepsilon})\xi_k^{\jmath}\Big]=0.$$

Likewise, detailed estimates reveal that

$$\begin{aligned}
&\lim_{\varepsilon\to 0}E\prod_{j_1=1}^{\kappa} h_{j_1}(Y^{\varepsilon,\imath}(t_{j_1}),\overline{\alpha}^{\varepsilon}(t_{j_1}))\\
&\quad\cdot\Big[\sum_{l\delta_\varepsilon=t}^{t+s}\overline{f}_{yy}(l\delta_\varepsilon,Y^{\varepsilon,\imath}_{ln_\varepsilon},\alpha^{\varepsilon}_{ln_\varepsilon})[Y^{\varepsilon,\imath}_{ln_\varepsilon+n_\varepsilon}-Y^{\varepsilon,\imath}_{ln_\varepsilon}]^2\Big]\\
&=\lim_{\varepsilon\to 0}E\prod_{j_1=1}^{\kappa} h_{j_1}(Y^{\varepsilon,\imath}(t_{j_1}),\overline{\alpha}^{\varepsilon}(t_{j_1}))\\
&\quad\cdot\Big[\sum_{l\delta_\varepsilon=t}^{t+s}\delta_\varepsilon f_{yy}(l\delta_\varepsilon,Y^{\varepsilon,\imath}_{ln_\varepsilon},\alpha^{\varepsilon}_{ln_\varepsilon})\frac{1}{n_\varepsilon}\\
&\quad\cdot\sum_{k=ln_\varepsilon}^{ln_\varepsilon+n_\varepsilon-1}\sum_{i=1}^{l}\sum_{j=1}^{m_i}E_k\Big[\sum_{\jmath=1}^{d}\sigma^{\imath\jmath}(\varepsilon k,\zeta_{ij})\xi_k^{\jmath}\Big]^2\nu_j^i I_{\{\alpha_k^\varepsilon\in\mathcal{M}_i\}}\Big]\\
&=E\prod_{j_1=1}^{\kappa} h_{j_1}(Y^{\imath}(t_{j_1}),\overline{\alpha}(t_{j_1}))\Big[\int_t^{t+s}f_{yy}(\tau,Y^{\imath}(\tau),\overline{\alpha}(\tau))\\
&\quad\cdot[\overline{\Sigma}(\tau,\overline{\alpha}(\tau))\overline{\Sigma}'(\tau,\overline{\alpha}(\tau))]^{\imath\imath}d\tau\Big],
\end{aligned}\tag{54}$$

where $\overline{\Sigma}(t,\alpha)$ and $[\overline{\Sigma}(t,\alpha)\overline{\Sigma}'(t,\alpha)]^{\imath\imath}$ are given by (20) and (23), respectively. In addition, similar to the estimates for (49), as $\varepsilon\to 0$,

$$\begin{aligned}
&E\prod_{j_1=1}^{\kappa} h_{j_1}(Y^{\varepsilon,\imath}(t_{j_1}),\overline{\alpha}^{\varepsilon}(t_{j_1}))\\
&\quad\cdot\sum_{l\delta_\varepsilon=t}^{t+s}[\overline{f}_y(l\delta_\varepsilon,Y^{\varepsilon,\imath}_{ln_\varepsilon},\alpha^{\varepsilon}_{ln_\varepsilon+n_\varepsilon})-\overline{f}_y(l\delta_\varepsilon,Y^{\varepsilon,\imath}_{ln_\varepsilon},\alpha^{\varepsilon}_{ln_\varepsilon})]\\
&\quad\cdot\sum_{k=ln_\varepsilon}^{ln_\varepsilon+n_\varepsilon-1}\varepsilon c^{\imath}(\varepsilon k,\alpha_k^{\varepsilon})\to 0,
\end{aligned}\tag{55}$$

$$\begin{aligned}
&E\prod_{j_1=1}^{\kappa} h_{j_1}(Y^{\varepsilon,\imath}(t_{j_1}),\overline{\alpha}^{\varepsilon}(t_{j_1}))\\
&\quad\cdot\sum_{l\delta_\varepsilon=t}^{t+s}[\overline{f}_{yy}(l\delta_\varepsilon,Y^{\varepsilon,\imath,+}_{ln_\varepsilon},\alpha^{\varepsilon}_{ln_\varepsilon+n_\varepsilon})-\overline{f}_{yy}(l\delta_\varepsilon,Y^{\varepsilon,\imath}_{ln_\varepsilon},\alpha^{\varepsilon}_{ln_\varepsilon})]\\
&\quad\cdot[Y^{\varepsilon,\imath}_{ln_\varepsilon+n_\varepsilon}-Y^{\varepsilon,\imath}_{ln_\varepsilon}]^2\to 0.
\end{aligned}\tag{56}$$

Similar to (55) and (56), as $\varepsilon\to 0$,

$$\begin{aligned}
&E\prod_{j_1=1}^{\kappa} h_{j_1}(Y^{\varepsilon,\imath}(t_{j_1}),\overline{\alpha}^{\varepsilon}(t_{j_1}))\\
&\quad\cdot\Big[\sum_{l\delta_\varepsilon=t}^{t+s}[\overline{f}_y(l\delta_\varepsilon,Y^{\varepsilon,\imath}_{ln_\varepsilon},\alpha^{\varepsilon}_{ln_\varepsilon+n_\varepsilon})-\overline{f}_y(l\delta_\varepsilon,Y^{\varepsilon,\imath}_{ln_\varepsilon},\alpha^{\varepsilon}_{ln_\varepsilon})]\\
&\quad\cdot\sum_{k=ln_\varepsilon}^{ln_\varepsilon+n_\varepsilon-1}\sqrt{\varepsilon}\sigma^{\imath\jmath}(\varepsilon k,\alpha_k^{\varepsilon})\xi_k^{\jmath}\Big]\to 0.
\end{aligned}\tag{57}$$

Combining the estimates obtained thus far and using (51) in conjunction with (47), we arrive at

$$\begin{aligned}
&E\prod_{j_1=1}^{\kappa} h_{j_1}(Y^{\varepsilon,\imath}(t_{j_1}),\overline{\alpha}^{\varepsilon}(t_{j_1}))\\
&\cdot\big(f(t+s,Y^{\varepsilon,\imath}(t+s),\overline{\alpha}^{\varepsilon}(t+s))-f(t,Y^{\varepsilon,\imath}(t),\overline{\alpha}^{\varepsilon}(t))\big)\\
&\quad\to E\prod_{j_1=1}^{\kappa} h_{j_1}(Y^{\imath}(t_{j_1}),\overline{\alpha}(t_{j_1}))\\
&\qquad\cdot\Big[\int_t^{t+s}\big[\big[\frac{\partial}{\partial\tau}+\mathcal{L}^{\imath}\big]f(\tau,Y^{\imath}(\tau),\overline{\alpha}(\tau))\\
&\qquad\quad+\overline{Q}f(\tau,Y^{\imath}(\tau),\cdot)(\overline{\alpha}(\tau))\big]d\tau\Big].
\end{aligned}\tag{58}$$

Equation (58) together with (46) then yields the desired assertion.

The same method works for the proof of $(S^{\varepsilon,0}(\cdot),\overline{\alpha}^{\varepsilon}(\cdot))$ to $(S^0(\cdot),\overline{\alpha}(\cdot))$. The proof is even simpler since no diffusion is involved. □

Acknowledgement. We thank the reviewers and the editors for their suggestions leading to much improvement. We also thank Wolfgang Runggaldier for commenting on an earlier draft of the paper.

REFERENCES

[1] G. Barone-Adesi and R. Whaley, Efficient analytic approximation of American option values, *J. Finance*, **42** (1987), 301-320.

[2] P. Billingsley, *Convergence of Probability Measures*, J. Wiley, New York, 1968.

[3] J. Buffington and R.J. Elliott, American Options with Regime Switching, to appear in *Theoretical Appl. Finance*.

[4] S. Chen, X. Li, and X.Y. Zhou, Stochastic linear quadratic regulators with indefinite control weight costs, *SIAM J. Control Optim.* **36** (1998), 1685-1702.

[5] G.B. Di Masi, Y.M. Kabanov and W.J. Runggaldier, Mean variance hedging of options on stocks with Markov volatility, *Theory Prob. Appl.* **39** (1994), 173-181.

[6] D. Duffie and H. Richardson, Mean-variance hedging in continuous time, *Ann. Appl. Probab.* **1** (1991), 1-15.

[7] E.J. Elton, M.J. Gruber, *Finance as a Dynamic Process*, Prentice Hall, Englewood Cliffs, 1975.

[8] J.C. Francis, *Investments: Analysis and Management*, McGraw-Hill, New York, 1976.

[9] R.R. Grauer and N.H. Hakansson, On the use of mean-variance and quadratic approximations in implementing dynamic investment strategies: a comparison of returns and investment policies, *Management Sci.* **39** (1993), 856-871.

[10] N.H. Hakansson, Multi-period mean-variance analysis: Toward a general theory of portfolio choice, *J. Finance*, **26** (1971), 857-884.

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

[12] H. J. Kushner, *Approximation and Weak Convergence Methods for Random Processes, with applications to Stochastic Systems Theory*, MIT Press, Cambridge, MA, 1984.

[13] D. Li and W.L. Ng, Optimal dynamic portfolio selection : Multi-period mean-variance formulation, *Math. Finance*, **10** (2000), 387-406.

[14] R.H. Liu, G. Yin, and Q. Zhang, Asymptotically optimal controls of hybrid linear quadratic regulators in discrete time, *Automatica*, **38** (2002), 409-419.

[15] D.G. Luenberger, *Investment Science*, Oxford University Press, New York, 1998.

[16] H. Markowitz, Portfolio selection, *J. Finance* **7** (1952), 77-91.

[17] H. Markowitz, *Portfolio Selection: Efficient Diversification of Investment*, John Wiley & Sons, New York, 1959.

[18] R.C. Merton, An intertemporal capital asset pricing model, *Econometrica*, **41** (1973), 867-888.

[19] J. Mossin, Optimal multiperiod portfolio policies, *J. Business* **41** (1968), 215-229.

[20] S. R. Pliska, *Introduction to Mathematical Finance: Discrete Time Models*, Blackwell, Oxford, 1997.
[21] P.A. Samuelson, Lifetime portfolio selection by dynamic stochastic programming, *Rev. Econ. Stat.* **51** (1969), 239-246.
[22] M. Schweizer, Approximation pricing and the variance-optimal martingale measure, *Ann. Probab.* **24** (1996), 206-236.
[23] G. Yin and Q. Zhang, *Continuous-Time Markov Chains and Applications: A Singular Perturbation Approach*, Springer-Verlag, New York, 1998.
[24] G. Yin and Q. Zhang, Singularly perturbed discrete-time Markov chains, *SIAM J. Appl. Math.* (2000) **61** (2000), 834-854.
[25] G. Yin, R.H. Liu, and Q. Zhang, Recursive algorithms for stock Liquidation: A stochastic optimization approach, *SIAM J. Optim.*, **13** (2002), 240-263.
[26] G. Yin, Q. Zhang, and G. Badowski, Asymptotic properties of a singularly perturbed Markov chain with inclusion of transient states, *Ann. Appl. Probab.* **10** (2000), 549-572.
[27] G. Yin, Q. Zhang, and G. Badowski, Discrete-time singularly perturbed Markov chains: Aggregation, occupation measures, and switching diffusion limit, *Adv. Appl. Probab.* **35** (2003), No. 2.
[28] J. Yong and X.Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer-Verlag, New York, 1999.
[29] T. Zariphopoulou, Investment-consumption models with transactions costs and Markov-chain parameters, *SIAM J. Control Optim.*, **30** (1992), 613-636.
[30] Q. Zhang, Stock trading: An optimal selling rule, *SIAM J. Control Optim.*, **40** (2001), 64-87.
[31] X.Y. Zhou and D. Li, Continuous-Time mean-variance portfolio selection: A stochastic LQ framework, *Appl. Math. Optim.* **42** (2000), 19-33.
[32] X.Y. Zhou and G. Yin, Markowitz's mean-variance portfolio selection with regime switching: A Continuous-time model, preprint, 2002.

PLACE PHOTO HERE

G. George Yin received his B.S. degree in mathematics, from the University of Delaware in 1983, M.S. in Electrical Engineering and Ph.D. in Applied Mathematics, from Brown University in 1987. Subsequently, he joined the Department of Mathematics, Wayne State University, where he became a professor in 1996. He severed on the editorial board of Stochastic Optimization & Design, the Mathematical Review Date Base Committee, and various conference program committees; he was the editor of SIAM Activity Group on Control and Systems Theory Newsletters, an Associate Editor of the IEEE Transactions on Automatic control, the SIAM Representative to the 34th CDC, and Chair of 1996 AMS-SIAM Summer Seminar in Applied Mathematics. He is also Chair of the 2003 AMS-IMA-SIAM Summer Research Conference "Mathematics of Finance."

PLACE PHOTO HERE

Xun Yu Zhou received his Ph. D. in Applied Mathematics in 1989 from Fudan University. During 1989-1991 and 1991-1993, he was doing postdoctoral research at Kobe University and University of Toronto, respectively. He was awarded Alexander von Humboldt Fellowship of Germany in 1991 and DAAD Research Fellowship of Germany in 1998. He has been with the Chinese University of Hong Kong since 1993 and is currently a Professor. As a principal investigator of numerous research grants, he has done research in optimal stochastic control, mathematical finance/insurance, and discrete-event manufacturing systems. He has served as Organizer and Program Committee Member for numerous international conferences, and has been invited to visit and/or give seminars at universities around the globe. He has published about 60 papers in refereed journals and his book entitled "Stochastic Controls: Hamiltonian Systems and HJB Equations" (coauthored with J. Yong) was published by Springer in 1999. He is a senior member of IEEE and a member of SIAM. He was on the editorial board of *IEEE Transactions on Automatic Control* from 1999 to 2002, and is currently on the editorial boards of *Operations Research* and *Mathematical Finance*.