

The State Equations Methods for Stochastic Control Problems

Lijin Wang^{1,2,*} and Fengshan Bai¹

¹ Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China.

² School of Mathematical Sciences, Graduate University of Chinese Academy of Sciences, Beijing 100049, China.

Received 28 February 2009; Accepted (in revised version) 7 July 2009

Available online 30 November 2009

Abstract. The state equations of stochastic control problems, which are controlled stochastic differential equations, are proposed to be discretized by the weak midpoint rule and predictor-corrector methods for the Markov chain approximation approach. Local consistency of the methods are proved. Numerical tests on a simplified Merton's portfolio model show better simulation to feedback control rules by these two methods, as compared with the weak Euler-Maruyama discretisation used by Krawczyk. This suggests a new approach of improving accuracy of approximating Markov chains for stochastic control problems.

AMS subject classifications: 93E20, 93E25, 90C39, 90C40

Key words: Stochastic optimal control, Markov chain approximation, Euler-Maruyama discretisation, midpoint rule, predictor-corrector methods, portfolio management.

1. Introduction

Many applications in finance and economics, such as the portfolio management problem, can be formulated as continuous time and continuous state stochastic control problems. These problems consists of the minimization of the cost function (in the finite horizon case)

$$J(\tau, x; u) = \mathbf{E} \left(\int_{\tau}^T L(x(t), u(x(t)), t) dt + s(x(T)) \middle| x(\tau) = x \right) \quad (1.1)$$

by choosing the optimal Markov feedback control policy $\hat{u}(x) \in \mathcal{U}(x) \subset \mathbb{R}^m$, for all states $x \in \mathcal{X} \subset \mathbb{R}^n$, subject to the state equation

$$dx(t) = f(x(t), u(x(t)), t) dt + g(x(t), u(x(t)), t) dW(t), \quad (1.2)$$

*Corresponding author. *Email addresses:* laeche1n@126.com (L. Wang), fbai@math.tsinghua.edu.cn (F. Bai)

which is a controlled stochastic differential equation (CSDE). Here $f : \mathbb{R}^n \times \mathbb{R}^m \times [0, +\infty] \mapsto \mathbb{R}^n$, $g : \mathbb{R}^n \times \mathbb{R}^m \times [0, +\infty] \mapsto \mathbb{R}^{n \times d}$ are continuous functions and satisfy conditions that are sufficient to have a solution for (1.2), and $W(t)$ is a d -dimensional Brownian motion. $\mathcal{U}(x)$ is called the set of admissible controls given x , and \mathcal{R} is the state space. The optimized cost function, denoted by $\hat{J}(\tau, x)$, is also called the optimal value function. That is,

$$\hat{J}(\tau, x) = \min_{u \in \mathcal{U}} J(\tau, x; u) = J(\tau, x; \hat{u}).$$

Except for some simple cases, e.g., those in [3, 4, 10], explicit solutions of continuous-time, continuous-state stochastic optimal control problems are very rare, and therefore numerical methods arise. The Markov chain approximation method is an efficient numerical approach which is widely used in financial economics. It goes back to Kushner (1977) and is described in Kushner and Dupuis [9]. The basic idea is to approximate the continuous-time, continuous-state problem by a discrete-time, discrete-state Markov chain model.

Krawczyk [7] proposed a totally different approach of finding transition probabilities based on discretisation of state equations, which is quite simple, intuitive and easy to understand. Effectiveness of his method is shown by [2, 7, 8]. The method of Krawczyk arouses our interest of investigating other numerical discretisations of the CSDE (1.2), as well as observing and comparing their effects on the final choice of the optimal control policy. As an example, we apply the midpoint rule instead of the Euler-Maruyama method to (1.2). The predictor-corrector methods (p-c^k) is also proposed in case of, e.g., nonlinear f or g .

Section 2 introduces the method of Krawczyk. The midpoint rule, p-c^k methods, and the Stratonovich stochastic differential equations, with which the midpoint rule is consistent, are stated in Section 3, where local consistency of Markov chains arising from the two kinds of methods are proved. Section 4 shows application of the two methods to a simplified Merton's portfolio model [7], and numerical experiments are performed in Section 5. Section 6 is a brief conclusion.

2. The method by Krawczyk

In the Markov chain approximation method, the state $x(t)$ is approximated by the Markov chain

$$\xi^h = \{\xi_k^h | k \in \mathbb{N}_0\} \subset \mathcal{R}_h, \quad (2.1)$$

where \mathcal{R}_h is the discrete state space, and $h = (h_1, \dots, h_n) \in \mathbb{R}_+^n$ is the state-step vector, and the cost function (1.1) is correspondingly changed to

$$J^{h, \Delta t^h}(\tau, x; u^h) = \mathbf{E} \left(\sum_{k=0}^{N-1} L(\xi_k^h, u_k^h, t_k^h) \Delta t_k^h + s(\xi_N^h) | \xi_0^h = x \right). \quad (2.2)$$

By choosing the Markov chain, its consistency with the state equation (1.2) is required. After the discrete model is constructed, the task then is to find the discrete optimal control $\hat{u}^h(x)$ for all $x \in \mathcal{R}_h$ to minimize the function $J^{h, \Delta t^h}$ in (2.2).

Denote $\hat{J}^{h,\Delta t^h}(\tau, x) = \inf_{u \in \mathcal{U}} J^{h,\Delta t^h}(\tau, x; u)$, it is deduced that

$$\hat{J}^{h,\Delta t^h}(\tau, x) = \min_{u \in \mathcal{U}} \left\{ L(x, u, \tau) \Delta t^h(x, u, \tau) + \sum_{y \in \mathcal{R}_h} p^{h,\Delta t^h}(x, y|u) \hat{J}^{h,\Delta t^h}(\tau + \Delta t^h(x, u, \tau), y) \right\},$$

with the boundary condition $\hat{J}^{h,\Delta t^h}(x, T) = s(x)$. This is called the dynamic programming equation (DPE), where $p^{h,\Delta t^h}(x, y|u)$ is the one-step transition probability from state x to y under the control u . \hat{u} and \hat{J} can be found by backward iteration algorithms on the DPE [9].

Classically, the determination of transition probabilities is based on the following partial differential equation, which is formally satisfied by the cost function $J(t, x; u)$:

$$J_t(t, x; u) + \mathcal{D}^{u(x)} J(t, x; u) + L(x, u(x)) = 0, \quad x \in \mathbb{R}^n, \quad t < T, \quad (2.3)$$

where

$$\mathcal{D}^u J(t, x) := J_x(t, x)^T f(x, u, t) + \frac{1}{2} \text{tr}(J_{xx}(t, x) g(x, u, t) g(x, u, t)^T),$$

and the boundary condition is $J(x, T; u) = s(x)$. Applying finite difference discretisation to the second order PDE (2.3), the transition probabilities from x to its neighboring states can be determined [9, 13].

Instead of Eq. (2.3), Krawczyk discretized the controlled stochastic differential equation (1.2) by using the weak Euler-Maruyama scheme. His method is briefly introduced below.

For the sake of simplicity, let the dimension of x and $W(t)$ equal 1. The Euler-Maruyama method applied to (1.2) gives

$$Y_{l+1} = Y_l + \delta f(Y_l, u_l, \tau_l) + g(Y_l, u_l, \tau_l) \Delta W_l, \quad (2.4)$$

where $\delta = T/N$ is the time discretisation step, $l = 0, 1, \dots, N-1$, $\Delta W_l = W(\tau_{l+1}) - W(\tau_l)$, and $u_l = u(x(\tau_l))$. Denote the discrete state space for time stage l by $\bar{\mathcal{R}}_l \subset \mathbb{R}^1$, and $\bar{U}_l = \max \bar{\mathcal{R}}_l$, $\bar{L}_l = \min \bar{\mathcal{R}}_l$. Then the discrete state space is $\{\bar{\mathcal{R}}_l\}_{l=0}^N$. Define the adjacency of states in the following way.

- Two states of $\bar{\mathcal{R}}_l$ are adjacent if no other state of $\bar{\mathcal{R}}_l$ lies between them.
- Given a point of the continuous state space, $x \in \mathcal{R}$, a pair of states, $\bar{x}^\ominus \in \bar{\mathcal{R}}_l$ and $\bar{x}^\oplus \in \bar{\mathcal{R}}_l$, is adjacent to x if the states are adjacent and $\bar{x}^\ominus < x < \bar{x}^\oplus$.
- Given $x \in \mathcal{R}$ with $x \geq \bar{U}_l$, define \bar{U}_l to be adjacent to x .
- Given $x \in \mathcal{R}$ with $x \leq \bar{L}_l$, define \bar{L}_l to be adjacent to x .
- Given $x \in \mathcal{R}$ with $x \in \bar{\mathcal{R}}_l$, define x to be adjacent to itself.

Suppose $Y_l = \bar{Y}_l \in \bar{\mathcal{R}}_l$ at stage l . Since $\Delta W_l = \zeta \sqrt{\delta}$, and $\zeta \sim \mathcal{N}(0, 1)$, simulating ζ with a two-point distributed random variable $\check{\zeta}$ with $P(\check{\zeta} = \pm 1) = \frac{1}{2}$, one gets a two-point distribution approximation $\Delta \tilde{W}_l$ for ΔW_l with

$$P(\Delta \tilde{W}_l = \pm \sqrt{\delta}) = \frac{1}{2}. \quad (2.5)$$

According to (2.4), given the control u_l , Y evolves from \bar{Y}_l to

$$Y_{l+1}^- = \bar{Y}_l + \delta f_l - g_l \sqrt{\delta} \quad \text{with prob. } \frac{1}{2}, \quad (2.6a)$$

$$Y_{l+1}^+ = \bar{Y}_l + \delta f_l + g_l \sqrt{\delta} \quad \text{with prob. } \frac{1}{2}, \quad (2.6b)$$

where

$$f_l = f(\bar{Y}_l, u_l, \tau_l), \quad g_l = g(\bar{Y}_l, u_l, \tau_l).$$

If $Y_{l+1}^- \notin \bar{\mathcal{R}}_{l+1}$ but there exist $\bar{Y}_{l+1}^{-\ominus} < \bar{Y}_{l+1}^{-\oplus}$ in $\bar{\mathcal{R}}_{l+1}$ adjacent to Y_{l+1}^- , then the transition probabilities are defined as

$$p(\bar{Y}_l, \bar{Y}_{l+1}^{-\oplus} | u_l) = \frac{1}{2} \frac{Y_{l+1}^- - \bar{Y}_{l+1}^{-\ominus}}{\bar{Y}_{l+1}^{-\oplus} - \bar{Y}_{l+1}^{-\ominus}}, \quad (2.7a)$$

$$p(\bar{Y}_l, \bar{Y}_{l+1}^{-\ominus} | u_l) = \frac{1}{2} \frac{\bar{Y}_{l+1}^{-\oplus} - Y_{l+1}^-}{\bar{Y}_{l+1}^{-\oplus} - \bar{Y}_{l+1}^{-\ominus}}. \quad (2.7b)$$

Similarly, if $Y_{l+1}^+ \notin \bar{\mathcal{R}}_{l+1}$ but there exist $\bar{Y}_{l+1}^{+\ominus} < \bar{Y}_{l+1}^{+\oplus}$ in $\bar{\mathcal{R}}_{l+1}$ adjacent to Y_{l+1}^+ , then

$$p(\bar{Y}_l, \bar{Y}_{l+1}^{+\oplus} | u_l) = \frac{1}{2} \frac{Y_{l+1}^+ - \bar{Y}_{l+1}^{+\ominus}}{\bar{Y}_{l+1}^{+\oplus} - \bar{Y}_{l+1}^{+\ominus}}, \quad (2.8a)$$

$$p(\bar{Y}_l, \bar{Y}_{l+1}^{+\ominus} | u_l) = \frac{1}{2} \frac{\bar{Y}_{l+1}^{+\oplus} - Y_{l+1}^+}{\bar{Y}_{l+1}^{+\oplus} - \bar{Y}_{l+1}^{+\ominus}}. \quad (2.8b)$$

If any of the states, for example $\bar{Y}_{l+1}^{+\ominus}$ and $\bar{Y}_{l+1}^{-\oplus}$ overlap each other, the respective probabilities must be summed up. This is the so-called *inverse distance method* of determining transition probabilities in [7].

3. The midpoint rule and p-c^k methods

By numerically solving stochastic differential equations, consistency of the numerical methods with the kind of stochastic differential equations is of essential importance, as shown by the example about inconsistency between the Heun methods and the Itô SDEs in [5].

In application, there are two kinds of frequently used stochastic differential equations, the Itô and Stratonovich equations. Due to their definitions, the solutions of the two kinds

of equations are usually different under the same drift and diffusion coefficient functions. However, they can transform to each other by adding or subtracting one term. Explicitly, the Itô SDE

$$dx = a(t, x)dt + b(t, x)dW(t) \quad (3.1)$$

is equivalent to its Stratonovich form

$$dx = a(t, x)dt + b(t, x) \circ dW(t) - \frac{1}{2} \frac{\partial b}{\partial x}(t, x)b(t, x)dt, \quad (3.2)$$

which means that (3.1) and (3.2) have the same solution [6, 11]. Writing a small 'o' before $dW(t)$ in (3.2) is a convention of denoting Stratonovich stochastic differential equations. In application, each of the two kinds of equations has its advantage, for example, the indefinite Itô integral $\int_0^t GdW$ is a martingale, while the ordinary differential chain rule holds for the Stratonovich differentials.

For SDEs of Stratonovich sense, the midpoint rule is proved to be convergent [12]. As an implicit method, the midpoint rule is usually stable and more accurate than the explicit Euler method. The cost, however, is that one must deal with the implicitness with fixed-point iteration in case of nonlinear coefficient functions. One kind of such iteration is the predictor-corrector approach. The predictor is usually an explicit method, which produces from x_n an initial approximation \tilde{x}_{n+1} to x_{n+1} , which is used as the initial point of iteration on the corrector. One calls, for example, a p-c^k method of Euler and midpoint rule to refer to a predictor-corrector method with k times of corrections, for which the Euler method is the predictor, and the midpoint rule the corrector.

In numerical time-discretisation of the CSDE (1.2), there is an additional term u to be treated. However, subject to the backward iteration algorithms used for finding the optimal control and value [7], the control u should always be taken as $u(x_n)$ by searching x_{n+1} from x_n , no matter what kind of numerical method is used.

We propose to apply the weak midpoint rule and p-c^k methods to (1.2), which is now assumed for simplicity to be of dimension 1. For the sake of consistency, (1.2) should be transformed to its Stratonovich form

$$dx(t) = \tilde{f}(x(t), u(x(t)), t)dt + g(x(t), u(x(t)), t) \circ dW(t), \quad (3.3)$$

where

$$\tilde{f}(x, u, t) = f(x, u, t) - \frac{1}{2}g(x, u, t)\frac{\partial g}{\partial x}(x, u, t). \quad (3.4)$$

The midpoint rule applied to (3.3) gives

$$x_{n+1} = x_n + \delta \tilde{f}\left(\frac{x_n + x_{n+1}}{2}, u_n, t_n\right) + \Delta W_n g\left(\frac{x_n + x_{n+1}}{2}, u_n, t_n\right), \quad (3.5)$$

where $\delta = t_{n+1} - t_n$, and $\Delta W_n = W(t_{n+1}) - W(t_n)$.

Simulating the noise by the two-point distribution (2.5), we get

$$x_{n+1}^- = x_n + \delta \tilde{f}\left(\frac{x_n + x_{n+1}^-}{2}, u_n, t_n\right) - \sqrt{\delta} g\left(\frac{x_n + x_{n+1}^-}{2}, u_n, t_n\right) \text{ with prob. } \frac{1}{2}, \quad (3.6)$$

$$x_{n+1}^+ = x_n + \delta \tilde{f}\left(\frac{x_n + x_{n+1}^+}{2}, u_n, t_n\right) + \sqrt{\delta} g\left(\frac{x_n + x_{n+1}^+}{2}, u_n, t_n\right) \text{ with prob. } \frac{1}{2}. \quad (3.7)$$

Finding the states that are adjacent to x_{n+1}^- and x_{n+1}^+ in the discrete state space $\bar{\mathcal{X}}_{n+1}$ of the approximating Markov chain, say, $\bar{x}_{n+1}^{-\ominus}$, $\bar{x}_{n+1}^{-\oplus}$, $\bar{x}_{n+1}^{+\ominus}$, $\bar{x}_{n+1}^{+\oplus}$, and determining the transition probabilities from x_n , assumed to lie in \mathcal{X}_n , to the four aforementioned states by the inverse distance method described in Section 2, we give a proof of the local consistency of the obtained Markov chain, which is not included in [7]. In the following, we write \bar{x}_n instead of x_n , since it is assumed to locate in the discrete state space of the n -th stage.

Theorem 3.1. *Suppose the coefficient functions \tilde{f} and g of the CSDE (3.3) are differentiable with respect to x , and continuous with respect to u and t . Then the approximating Markov chain arising from the weak midpoint rule (3.6)-(3.7) with transition probabilities*

$$p(\bar{x}_n, \bar{x}_{n+1}^{-\oplus} | u_n) = \frac{1}{2} \frac{x_{n+1}^- - \bar{x}_{n+1}^{-\ominus}}{\bar{x}_{n+1}^{-\oplus} - \bar{x}_{n+1}^{-\ominus}}, \quad (3.8)$$

$$p(\bar{x}_n, \bar{x}_{n+1}^{-\ominus} | u_n) = \frac{1}{2} \frac{\bar{x}_{n+1}^{-\oplus} - x_{n+1}^-}{\bar{x}_{n+1}^{-\oplus} - \bar{x}_{n+1}^{-\ominus}}, \quad (3.9)$$

$$p(\bar{x}_n, \bar{x}_{n+1}^{+\oplus} | u_n) = \frac{1}{2} \frac{x_{n+1}^+ - \bar{x}_{n+1}^{+\ominus}}{\bar{x}_{n+1}^{+\oplus} - \bar{x}_{n+1}^{+\ominus}}, \quad (3.10)$$

$$p(\bar{x}_n, \bar{x}_{n+1}^{+\ominus} | u_n) = \frac{1}{2} \frac{\bar{x}_{n+1}^{+\oplus} - x_{n+1}^+}{\bar{x}_{n+1}^{+\oplus} - \bar{x}_{n+1}^{+\ominus}}, \quad (3.11)$$

and $p(\bar{x}_n, \bar{y}_{n+1} | u_n) = 0$ for all $\bar{y}_{n+1} \in \bar{\mathcal{X}}_{n+1}$ but not in $\{\bar{x}_{n+1}^{-\ominus}, \bar{x}_{n+1}^{-\oplus}, \bar{x}_{n+1}^{+\ominus}, \bar{x}_{n+1}^{+\oplus}\}$, is locally consistent.

Proof. From the scheme (3.6)-(3.7), it follows that

$$k_{1,n} := x_{n+1}^- - \bar{x}_n = \delta \tilde{f}\left(\frac{\bar{x}_n + x_{n+1}^-}{2}, u_n, t_n\right) - \sqrt{\delta} g\left(\frac{\bar{x}_n + x_{n+1}^-}{2}, u_n, t_n\right), \quad (3.12)$$

$$k_{2,n} := x_{n+1}^+ - \bar{x}_n = \delta \tilde{f}\left(\frac{\bar{x}_n + x_{n+1}^+}{2}, u_n, t_n\right) + \sqrt{\delta} g\left(\frac{\bar{x}_n + x_{n+1}^+}{2}, u_n, t_n\right). \quad (3.13)$$

Denote

$$\begin{aligned} h_n^- &= \bar{x}_{n+1}^{-\oplus} - \bar{x}_{n+1}^{-\ominus}, & h_n^+ &= \bar{x}_{n+1}^{+\oplus} - \bar{x}_{n+1}^{+\ominus}, & \alpha_n &= x_{n+1}^- - \bar{x}_{n+1}^{-\ominus}, \\ \beta_n &= \bar{x}_{n+1}^{-\oplus} - x_{n+1}^-, & \mu_n &= x_{n+1}^+ - \bar{x}_{n+1}^{+\ominus}, & \nu_n &= \bar{x}_{n+1}^{+\oplus} - x_{n+1}^+, \end{aligned}$$

then

$$\alpha_n + \beta_n = h_n^-, \quad \mu_n + \nu_n = h_n^+.$$

Let $\Delta\bar{X}_n = \bar{X}_{n+1} - \bar{X}_n$, where the capital X represents the random state variable. We have

$$\begin{aligned} & \mathbf{E}(\Delta\bar{X}_n | u_n, \bar{X}_n = \bar{x}_n) \\ &= \frac{1}{2} \frac{\beta_n}{h_n^-} (-\alpha_n + k_{1,n}) + \frac{1}{2} \frac{\alpha_n}{h_n^-} (\beta_n + k_{1,n}) + \frac{1}{2} \frac{\nu_n}{h_n^+} (-\mu_n + k_{2,n}) + \frac{1}{2} \frac{\mu_n}{h_n^+} (\nu_n + k_{2,n}), \\ &= \frac{1}{2} (k_{1,n} + k_{2,n}). \end{aligned} \quad (3.14)$$

Denote

$$\begin{aligned} \Delta x_n^- &= \frac{\bar{x}_n + x_{n+1}^-}{2} - \bar{x}_n = \frac{x_{n+1}^- - \bar{x}_n}{2}, & \Delta x_n^+ &= \frac{\bar{x}_n + x_{n+1}^+}{2} - \bar{x}_n = \frac{x_{n+1}^+ - \bar{x}_n}{2}, \\ \tilde{f}_n &= \tilde{f}(\bar{x}_n, u_n, t_n), & g_n &= g(\bar{x}_n, u_n, t_n). \end{aligned} \quad (3.15)$$

Consequently, we have

$$\begin{aligned} & k_{1,n} + k_{2,n} \\ &= \delta \left[\tilde{f} \left(\frac{\bar{x}_n + x_{n+1}^-}{2}, u_n, t_n \right) + \tilde{f} \left(\frac{\bar{x}_n + x_{n+1}^+}{2}, u_n, t_n \right) \right] \\ & \quad + \sqrt{\delta} \left[g \left(\frac{\bar{x}_n + x_{n+1}^+}{2}, u_n, t_n \right) - g \left(\frac{\bar{x}_n + x_{n+1}^-}{2}, u_n, t_n \right) \right] \\ &= \delta \left[\tilde{f}_n + \frac{\partial \tilde{f}}{\partial x}(\bar{x}_n, u_n, t_n) \Delta x_n^- + o(\Delta x_n^-) + \tilde{f}_n + \frac{\partial \tilde{f}}{\partial x}(\bar{x}_n, u_n, t_n) \Delta x_n^+ + o(\Delta x_n^+) \right] \\ & \quad + \sqrt{\delta} \left[g_n + \frac{\partial g}{\partial x}(\bar{x}_n, u_n, t_n) \Delta x_n^+ + o(\Delta x_n^+) - g_n - \frac{\partial g}{\partial x}(\bar{x}_n, u_n, t_n) \Delta x_n^- - o(\Delta x_n^-) \right] \\ &= \delta \left[2\tilde{f}_n + \frac{\partial \tilde{f}}{\partial x}(\bar{x}_n, u_n, t_n) (\Delta x_n^- + \Delta x_n^+) + o(\sqrt{\delta}) \right] \\ & \quad + \sqrt{\delta} \left[\frac{\partial g}{\partial x}(\bar{x}_n, u_n, t_n) (\Delta x_n^+ - \Delta x_n^-) + o(\sqrt{\delta}) \right]. \end{aligned} \quad (3.16)$$

It is clear that $\Delta x_n^- + \Delta x_n^+ = \mathcal{O}(\delta^\gamma)$ for some $\gamma \geq \frac{1}{2}$, as $\delta \rightarrow 0$. Moreover,

$$\begin{aligned} \Delta x_n^+ - \Delta x_n^- &= \frac{1}{2} \delta \left[\frac{\partial \tilde{f}}{\partial x}(\bar{x}_n, u_n, t_n) (\Delta x_n^+ - \Delta x_n^-) + o(\sqrt{\delta}) \right] \\ & \quad + \frac{1}{2} \sqrt{\delta} \left[2g_n + \frac{\partial g}{\partial x}(\bar{x}_n, u_n, t_n) (\Delta x_n^+ - \Delta x_n^-) + o(\sqrt{\delta}) \right] \\ &= \sqrt{\delta} g_n + o(\delta). \end{aligned} \quad (3.17)$$

Substituting (3.17) into (3.16), we obtain

$$k_{1,n} + k_{2,n} = 2\delta \tilde{f}_n + \delta g_n \frac{\partial g}{\partial x}(\bar{x}_n, u_n, t_n) + o(\delta), \quad (3.18)$$

which together with (3.14) implies that

$$\begin{aligned}\mathbf{E}(\Delta\bar{X}_n|u_n, \bar{X}_n = \bar{x}_n) &= \delta\tilde{f}_n + \frac{1}{2}\delta g_n \frac{\partial g}{\partial x}(\bar{x}_n, u_n, t_n) + o(\delta) \\ &= \delta f_n + o(\delta),\end{aligned}\quad (3.19)$$

where $f_n = f(\bar{x}_n, u_{n+1}, t_n)$. The result (3.19) is the first condition of the local consistency. Further,

$$\begin{aligned}&\mathbf{E}(|\Delta\bar{X}_n - \mathbf{E}(\Delta\bar{X}_n|u_n, \bar{X}_n = \bar{x}_n)|^2|u_n, \bar{X}_n = \bar{x}_n) \\ &= \mathbf{E}(|\Delta\bar{X}_n|^2|u_n, \bar{X}_n = \bar{x}_n) - (\mathbf{E}(\Delta\bar{X}_n|u_n, \bar{X}_n = \bar{x}_n))^2 \\ &= \bar{\beta}_n(-\alpha_n + k_{1,n})^2 + \bar{\alpha}_n(\beta_n + k_{1,n})^2 + \bar{v}_n(-\mu_n + k_{2,n})^2 \\ &\quad + \bar{\mu}_n(\nu_n + k_{2,n})^2 - \frac{1}{4}(k_{1,n} + k_{2,n})^2 \\ &= \frac{1}{4}(k_{1,n}^2 + k_{2,n}^2) - \frac{1}{2}k_{1,n}k_{2,n} + \frac{1}{2}(\alpha_n\beta_n + \mu_n\nu_n),\end{aligned}\quad (3.20)$$

where

$$\bar{\alpha}_n = \frac{1}{2}\frac{\alpha_n}{h_n^-}, \quad \bar{\beta}_n = \frac{1}{2}\frac{\beta_n}{h_n^-}, \quad \bar{\mu}_n = \frac{1}{2}\frac{\mu_n}{h_n^+}, \quad \bar{\nu}_n = \frac{1}{2}\frac{\nu_n}{h_n^+},\quad (3.21)$$

and the derivation of the last step of (3.20) is based on the relations

$$\bar{\alpha}_n + \bar{\beta}_n = \bar{\mu}_n + \bar{\nu}_n = \frac{1}{2}, \quad \bar{\alpha}_n\beta_n = \alpha_n\bar{\beta}_n, \quad \bar{\mu}_n\nu_n = \mu_n\bar{\nu}_n.\quad (3.22)$$

Similar to the derivation of (3.16), one can show that

$$\begin{aligned}k_{1,n}^2 + k_{2,n}^2 &= \delta \left[\left(g_n + \frac{\partial g}{\partial x}(\bar{x}_n, u_n, t_n)\Delta x_n^- + o(\Delta x_n^-) \right)^2 \right. \\ &\quad \left. + \left(g_n + \frac{\partial g}{\partial x}(\bar{x}_n, u_n, t_n)\Delta x_n^+ + o(\Delta x_n^+) \right)^2 \right] + o(\delta) \\ &= 2\delta g_n^2 + o(\delta),\end{aligned}\quad (3.23)$$

and

$$k_{1,n}k_{2,n} = -\delta g_n^2 + o(\delta).\quad (3.24)$$

Substituting (3.23)-(3.24) into (3.20), and noticing that

$$\frac{1}{2}(\alpha_n\beta_n + \mu_n\nu_n) \leq \frac{1}{2} \left(\frac{(\alpha_n + \beta_n)^2}{2} + \frac{(\mu_n + \nu_n)^2}{2} \right) = \frac{(h_n^-)^2 + (h_n^+)^2}{4},$$

we have

$$\mathbf{E} \left(|\Delta\bar{X}_n - \mathbf{E}(\Delta\bar{X}_n|u_n, \bar{X}_n = \bar{x}_n)|^2 | u_n, \bar{X}_n = \bar{x}_n \right) = \delta g_n^2 + o(\delta) + o(\|h_n\|_2),\quad (3.25)$$

where h_n is the vector of state steps at stage n . (3.25) is the second condition of the local consistency [13].

Denote $h := \sup_n \{\|h_n\|_\infty\}$. The third condition requires that $\sup_{n,\omega} |\Delta \bar{X}_n| \rightarrow 0$, as $h \rightarrow 0$. For this, we observe

$$|-\alpha_n + k_{1,n}|, |\beta_n + k_{1,n}|, |-\mu_n + k_{2,n}|, |\nu_n + k_{2,n}|.$$

For any n , these are four realizations of the distance with respect to the realizations of the random sample point ω . Under the continuity of the functions \tilde{f} and g , as well as the boundedness of the state space, the admissible control space and the time interval, \tilde{f} and g must be bounded. Suppose $|\tilde{f}| \leq M$ and $|g| \leq N$, for all x, u and t . Then for any n , the supremum of the four distances must be smaller than

$$h + \delta M + \sqrt{\delta} N,$$

which tends to zero as $h \rightarrow 0$, since it is assumed that $\delta \rightarrow 0$ as $h \rightarrow 0$ in the construction of the Markov chain [9].

For the trivial case that $\alpha_n = \beta_n = 0$ or $\mu_n = \nu_n = 0$, or both happen, the proof follows the same line and the conclusion remains invariant. \square

As mentioned above, however, with all the benefit of the implicit midpoint rule, it is difficult to implement in case of nonlinear f or g . The predictor-corrector methods supply a compensation in such case. For the CSDE (3.3), the p-c^k ($k \geq 1$) method of Euler-Maruyama and midpoint rule produces the scheme

$$\tilde{x}_{n+1} = x_n + \delta f(x_n, u_n, t_n) + \Delta W_n g(x_n, u_n, t_n), \quad (3.26)$$

$$x_{n+1}^{(1)} = x_n + \delta \tilde{f}\left(\frac{x_n + \tilde{x}_{n+1}}{2}, u_n, t_n\right) + \Delta W_n g\left(\frac{x_n + \tilde{x}_{n+1}}{2}, u_n, t_n\right),$$

\vdots

$$x_{n+1}^{(k)} = x_n + \delta \tilde{f}\left(\frac{x_n + x_{n+1}^{(k-1)}}{2}, u_n, t_n\right) + \Delta W_n g\left(\frac{x_n + x_{n+1}^{(k-1)}}{2}, u_n, t_n\right), \quad (3.27)$$

where the upper index (i) refers to the i -th correction. Note that the predictor, the Euler-Maruyama method is applied to the Itô equation (1.2), while the corrector, the midpoint rule, is applied k times to the Stratonovich equation (3.3).

Again, simulating the noise ΔW_n by the two-point distribution in (2.5), we can prove the local consistency of the Markov chain resulting from the weak p-c^k method of Euler-Maruyama and midpoint rule.

Theorem 3.2. *Under the same assumption of Theorem 3.1, the Markov chain arising from the weak p-c^k ($k \geq 1$) method based on (3.26)-(3.27), with predictor the Euler-Maruyama method, and corrector the midpoint rule, is locally consistent for all $k \geq 1$, where the transition probabilities are the same with that given in Theorem 3.1.*

Proof. The proof is analogous to that of Theorem 3.1, so we omit it. \square

Effectiveness of the two kinds of methods given above are tested on a portfolio model described in the following section. Comparison between the two methods and the Euler-Maruyama method is observed.

4. Applications to simplified Merton's portfolio model

In [3], a simplified version of Merton's [10] optimal portfolio selection model is studied and analytically solved. It is used as a test model for numerical methods in [7].

In the model, the aim of the portfolio selection is to maximize the expected discounted utility

$$J(0, x(0); u) = \mathbf{E} \left(\int_0^T e^{-\rho t} [U_2(t)]^\gamma dt | x(0) = x_0 \right) \quad (4.1)$$

with given discount rate $\rho > 0$, where $[U_2(t)]^\gamma$ is the agent's utility function. The constraint equation is a controlled stochastic differential equation describing evolution of the wealth $x(t)$:

$$dx = (1 - u_1)rx dt + u_1x(\alpha dt + \sigma dW) - U_2 dt, \quad (4.2)$$

where r, α, σ are constants with $r < \alpha, \sigma > 0$, and $W(t)$ is a one-dimensional standard Wiener process. $u_1(t)$ is the fraction of the wealth invested in the risky asset at time t , and $U_2(t)$ denotes the consumption rate.

The optimal two-dimensional control $\hat{u} = [\hat{u}_1(x), \hat{U}_2(x)]$, which satisfies the condition

$$0 \leq u_1 \leq 1, \quad U_2 \geq 0, \quad (4.3)$$

are to be found that maximize the expected discounted utility (4.1).

The solution of this problem is explicit. The optimal control is

$$\hat{u}_1 = \frac{\alpha - r}{\sigma^2(1 - \gamma)}, \quad \hat{U}_2(\tau, x) = [e^{\rho\tau} g(\tau)]^{\frac{1}{\gamma-1}} x, \quad (4.4)$$

where

$$g(\tau) = e^{-\rho\tau} \left[\frac{1 - \gamma}{\rho - \nu\gamma} \left(1 - e^{-\frac{(\rho - \nu\gamma)}{1 - \gamma}(T - \tau)} \right) \right]^{1 - \gamma} \quad (4.5)$$

with

$$\nu = \frac{(\alpha - r)^2}{2\sigma^2(1 - \gamma)} + r. \quad (4.6)$$

The optimal value function is

$$H(\tau, x) = g(\tau)x^\gamma. \quad (4.7)$$

Krawczyk used the so-called re-scaling technique in [7], that is, to replace U_2 by u_2x , so that the components of the control vector $u = [u_1, u_2]$ remain in comparable magnitude, which is much better for efficiency of numerical methods.

For the state equation (4.2), the Euler-Maruyama discretisation gives

$$x_{n+1} = (1 + \delta m + u_1\sigma\Delta W_n)x_n, \quad (4.8)$$

where $m = (1 - u_1)r + u_1\alpha - u_2$.

To use the midpoint rule, we transform the Itô equation (4.2) to its Stratonovich form

$$dx = (1 - u_1)rxdt + u_1x(\alpha dt + \sigma \circ dW) - u_2xdt - \frac{1}{2}u_1^2\sigma^2xdt. \quad (4.9)$$

The midpoint rule applied to (4.9) gives

$$\begin{aligned} & \left(1 - \frac{\delta}{2} \left(m - \frac{u_1^2\sigma^2}{2}\right) - \frac{u_1\sigma\Delta W_n}{2}\right)x_{n+1} \\ &= \left(1 + \frac{\delta}{2} \left(m - \frac{u_1^2\sigma^2}{2}\right) + \frac{u_1\sigma\Delta W_n}{2}\right)x_n. \end{aligned} \quad (4.10)$$

Taking the x_{n+1} given by (4.8) as the predicted value \tilde{x}_{n+1} , we have the p-c^k discretisation with Euler-Maruyama and midpoint rule for (4.9)

$$x_{n+1}^{(1)} = x_n + \delta \left(m - \frac{u_1^2\sigma^2}{2}\right) \frac{x_n + \tilde{x}_{n+1}}{2} + \Delta W_n u_1 \sigma \frac{x_n + \tilde{x}_{n+1}}{2}, \quad (4.11)$$

⋮

$$x_{n+1}^{(k)} = x_n + \delta \left(m - \frac{u_1^2\sigma^2}{2}\right) \frac{x_n + x_{n+1}^{(k-1)}}{2} + \Delta W_n u_1 \sigma \frac{x_n + x_{n+1}^{(k-1)}}{2}. \quad (4.12)$$

In our implementation of the midpoint rule (4.10) and the p-c^k method (4.11)-(4.12), we also use the constraints given in [7]:

$$x_n(1 + \delta(r + u_{1,n}(\alpha - r) - u_{2,n})) \geq 0, \quad (4.13a)$$

$$0 \leq u_{1,n} \leq 1, \quad u_{2,n} \geq 0, \quad (4.13b)$$

for all $n \geq 0$, and ΔW_n is approximated by the two-point distribution random variable $\Delta \tilde{W}_n$ described in (2.5).

In the next section, numerical experiments are performed on this portfolio model.

5. Numerical experiments

Krawczyk's Markov chain approximating method was applied to the simplified Merton's model (4.1)-(4.3) in [7]. Later in [1], a suite of MATLAB[®] functions implementing this method, named **SOCSol41**, was described. In this package, the computation of the optimal control is done by the file SOCSol.m, the visualization of the optimal control by ContRule.m, and that of the optimal value by ValGraph.m. For our numerical tests about the midpoint rule and the p-c^k methods, some parts of the package relating to the discretisation of the state equations, e.g., the DeltaFunctionFile, CostStoch.m and GenSim.m must be modified accordingly.

While the Euler-Maruyama discretisation method does work in finding the optimal control $[\hat{u}_1, \hat{u}_2]$, as reported in [7], the convergence of u_2 seems not so satisfying, especially

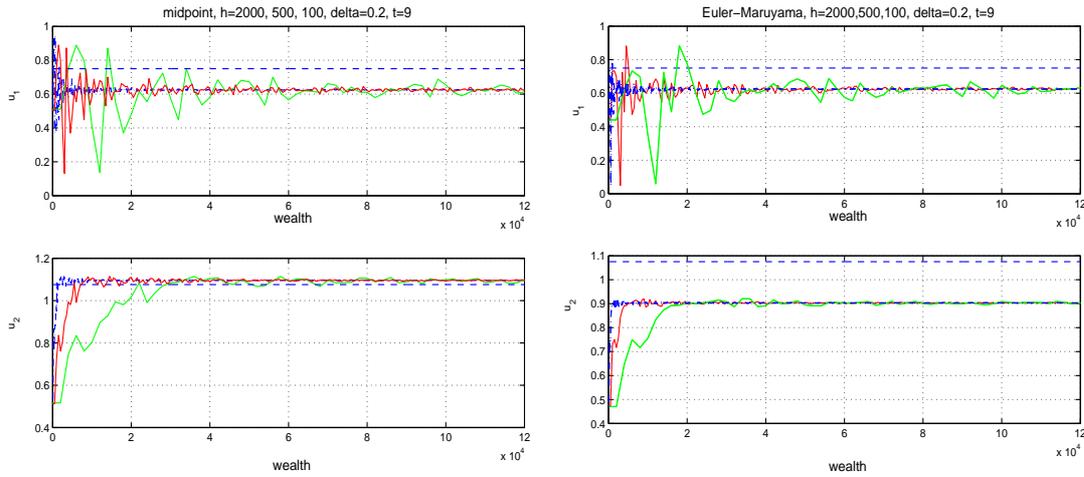


Figure 1: Control rules created by midpoint rule as $\delta = 0.2$ (left) and Euler-Maruyama method as $\delta = 0.2$ (right).

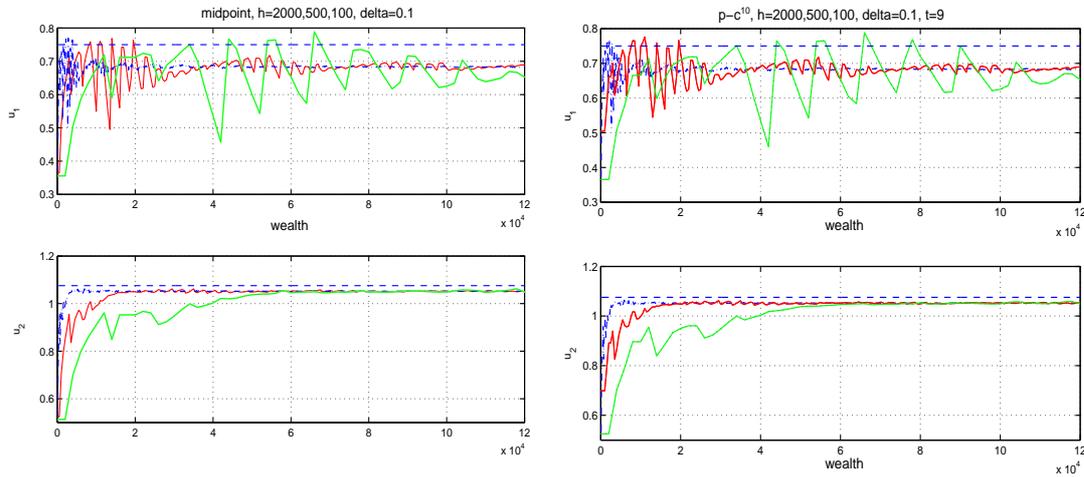


Figure 2: Control rules created by midpoint rule as $\delta = 0.1$ (left) and p-c¹⁰ method as $\delta = 0.1$ (right).

at time $t = 9$. It requires very small time-step, say, $\delta = 0.02$. This is illustrated in Figs. 1 (right), 3 and 4 (right). It can be seen that, there is an obvious gap between the theoretical and numerical \hat{u}_2 . Even at $\delta = 0.05$, the gap is approximately 0.1, which means that the relative error is nearly 9 percent.

In the following, the optimal control and value arising from the midpoint rule (4.10) and the p-c¹⁰ method of Euler-Maruyama and midpoint rule (4.12) are observed, wherein better convergence of u_2 is seen.

The parameters for the portfolio model described in (4.1)-(4.3) take the same values as in [7], i.e., $\alpha = 0.11$, $\sigma = 0.4$, $r = 0.05$, $\rho = 0.11$, $\gamma = \frac{1}{2}$. The initial wealth is $x_0 = \$100,000$, and the time interval to observe is $T = 10$ years. The optimal controls are

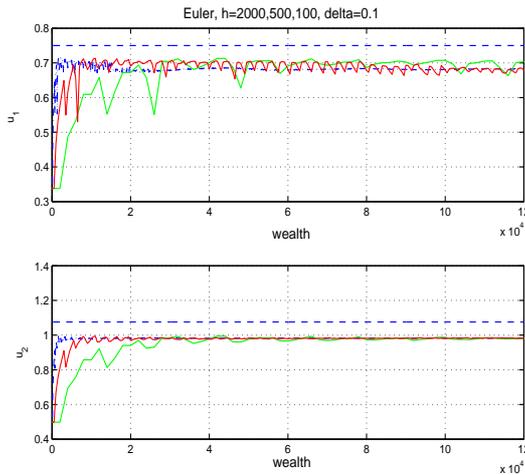


Figure 3: Control rules created by Euler-Maruyama method as $\delta = 0.1$.

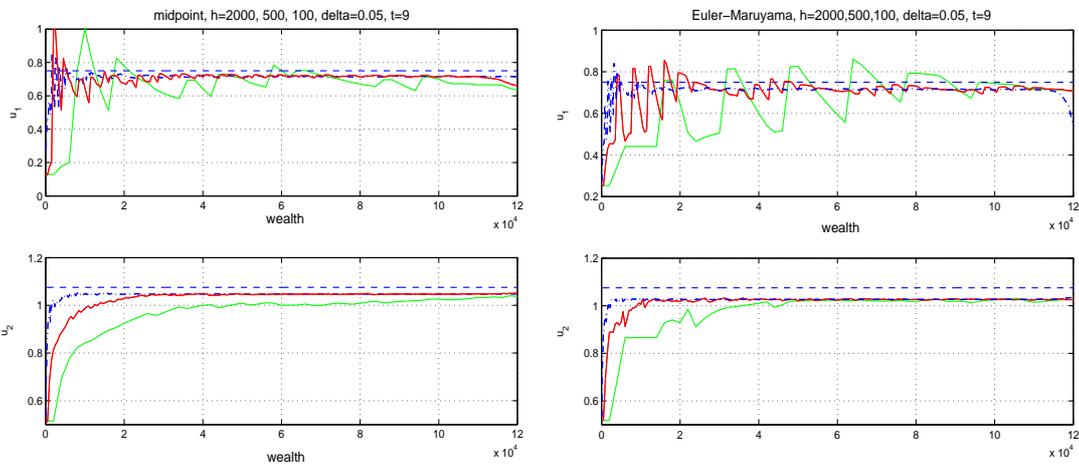


Figure 4: Control rules created by midpoint rule as $\delta = 0.05$ (left) and Euler-Maruyama method as $\delta = 0.05$ (right).

computed at time $t = 9$, which is a 'problem-time' for the Euler-Maruyama discretisation, while the optimal value is observed with respect to the initial ($t = 0$) wealth.

The left of Figs. 1, 2 and 4 is obtained by using the midpoint rule, while the right of Figs. 1 and 4, as well as Fig. 3 are given by using the Euler-Maruyama discretisation. The right of Fig. 2 is produced by the p-c¹⁰ method of the Euler-Maruyama and midpoint rule.

The time step size is $\delta = 0.2$ for Fig. 1, $\delta = 0.1$ for Figs. 2-3, and $\delta = 0.05$ for Fig. 4. In each figure, the state step size h takes the values 2000 (green solid line), 500 (red solid line) and 100 (blue dash-dot line), and the dashed lines therein are the theoretical optimal controls $\hat{u}_1 = 0.75$ and $\hat{u}_2 = 1.0756$ from (4.4)-(4.6). The sub-figures above are for u_1 , and below for u_2 .

The convergence of u_1 to the theoretical value 0.75 as δ and h decrease are obvious and comparable for the three methods. However, the situation for u_2 is different. By the

Table 1: Numerical results of $e = \|u - \hat{u}\|_2$ with different values of wealth.

Wealth	Midpoint Rule		Euler-Maruyama	
	$\delta = 0.1$	$\delta = 0.05$	$\delta = 0.1$	$\delta = 0.05$
20,000	0.0757	0.0402	0.1172	0.0636
40,000	0.0676	0.0413	0.1157	0.0595
60,000	0.0703	0.0420	0.1176	0.0610
80,000	0.0728	0.0455	0.1145	0.0594
100,000	0.0690	0.0448	0.1165	0.0608

Table 2: CPU comparison for different numerical methods.

	Midpoint Rule (seconds)	Euler-Maruyama (seconds)	p-c ¹⁰ (seconds)
$h = 2000$	301.547	285.781	309.015
$h = 500$	809.407	818.640	875.640
$h = 100$	2517.094	2157.265	2268.532

midpoint rule, convergence of u_2 can already be seen at $\delta = 0.2$, while distance between theoretical and numerical u_2 is still obvious at $\delta = 0.05$ by the Euler-Maruyama method. The right of Fig. 2 shows that the p-c¹⁰ method behaves also well in simulating u_2 .

To consider the approximation of u_1 and u_2 synthetically, that is, to observe the value $e = \sqrt{(u_1 - \hat{u}_1)^2 + (u_2 - \hat{u}_2)^2}$, which is the l^2 -norm of the error vector $u - \hat{u}$, the following figures illustrating the error e by using the three aforementioned methods are drawn against the wealth x at $t = 9$.

In left of Fig. 5, the error curves arising from the midpoint rule (red solid line) and the p-c¹⁰ method (blue dash-dot line) coincide almost, while that from the Euler-Maruyama method (blue dashed line) locates up them. The right of Fig. 5 illustrates the errors from the midpoint rule (red solid line) and the Euler-Maruyama method (blue dashed line), which locate lower and closer than those in the left figure of Fig. 5. This is owing to the reduction of the time step δ in computing u , which takes the value 0.1 for the left while 0.05 for the right of Fig. 5. The state step for computing u is $h = 100$ for both figures. Concrete numerical results of $e = \|u - \hat{u}\|_2$ are listed against certain wealth values in Table 1, i.e. for $x = 20000, 40000, 60000, 80000$, and 100000 .

Fig. 5 shows that the midpoint rule and p-c¹⁰ method are more accurate than the Euler-Maruyama method, while the digits in Table 1 show decline of the errors with decrease of δ , by the two mentioned methods.

The CPU time needed for computing the optimal control rules by the three methods is compared in Table 2. For this we have taken $\delta = 0.1$, $t = 9$, and $x \in [0, 120000]$.

The elapsed time can be slightly affected by different programming strategies, e.g., direct formulating or taking a long formula into two shorter parts. Nevertheless, it is clear that the consumed CPU time are comparable by the midpoint rule, the p-c¹⁰ methods, and the Euler-Maruyama method.

The effect of the three methods in simulating the optimal value is also tested, for which the MATLAB file ValGraph.m uses the control rule computed by SOCSol.m to compute the

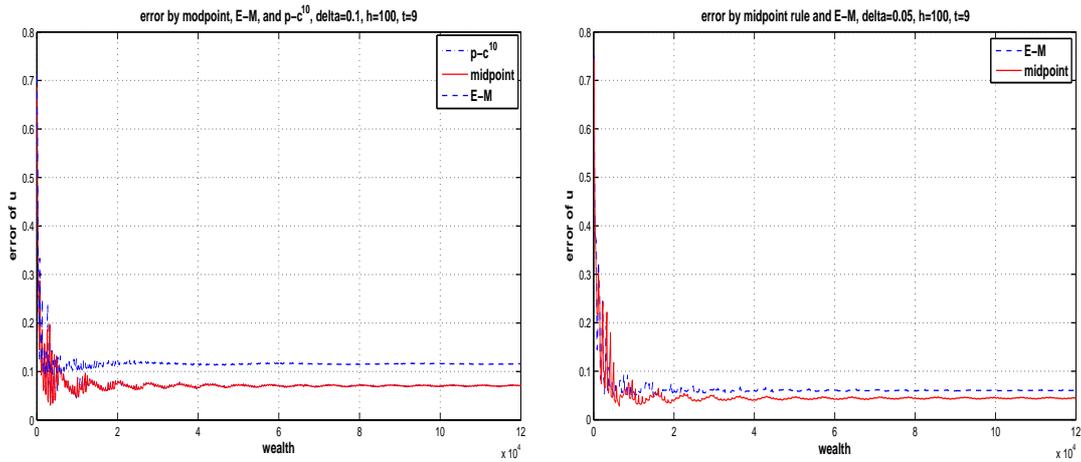


Figure 5: Error in u created by the three methods as $\delta = 0.1$ (left) and by midpoint and E-M methods as $\delta = 0.05$ (right).

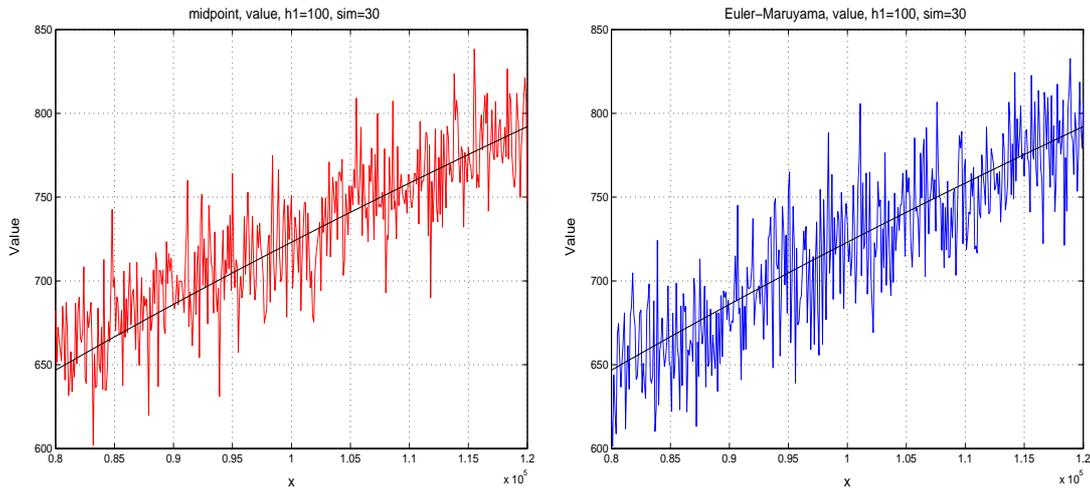


Figure 6: Expected value created by the midpoint rule in 30 simulations (left) and the E-M method in 30 simulations (right).

expected values. For the following figures, the controls are computed at $h = 100$ and $\delta = 0.1$.

In Fig. 6, the black line stretches in the middle is the theoretical optimal value which is calculated from (4.7). The oscillating curves in Fig. 6 show the expected optimal value arising from the midpoint rule and the Euler-Maruyama method, respectively, which are resulted from 30 simulations. The step used along the x -axis by plotting the two pictures is $h1 = 100$, which is relatively fine for the much larger observation interval $x \in [80000, 120000]$. This fact, together with the less amount of simulations imply the violent oscillation of the curves.

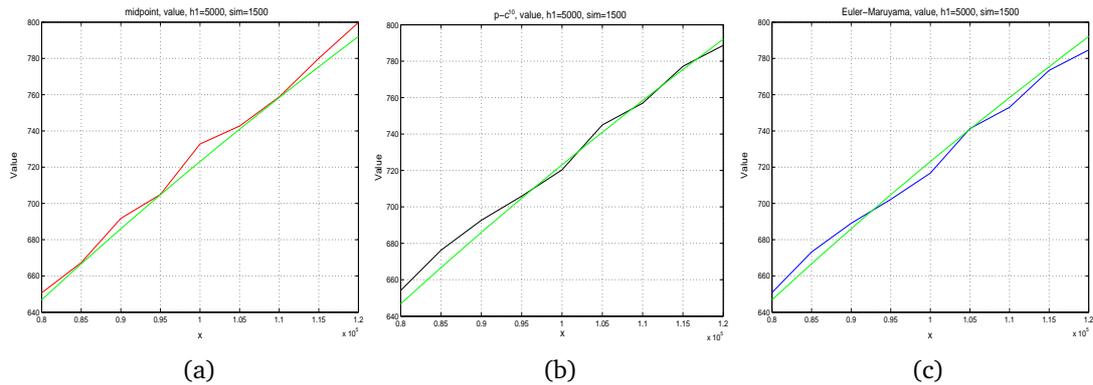


Figure 7: Expected value given by (a) the midpoint rule; (b) the $p-c^{10}$ method; and (c) the Euler-Maruyama method, with 1500 simulations.

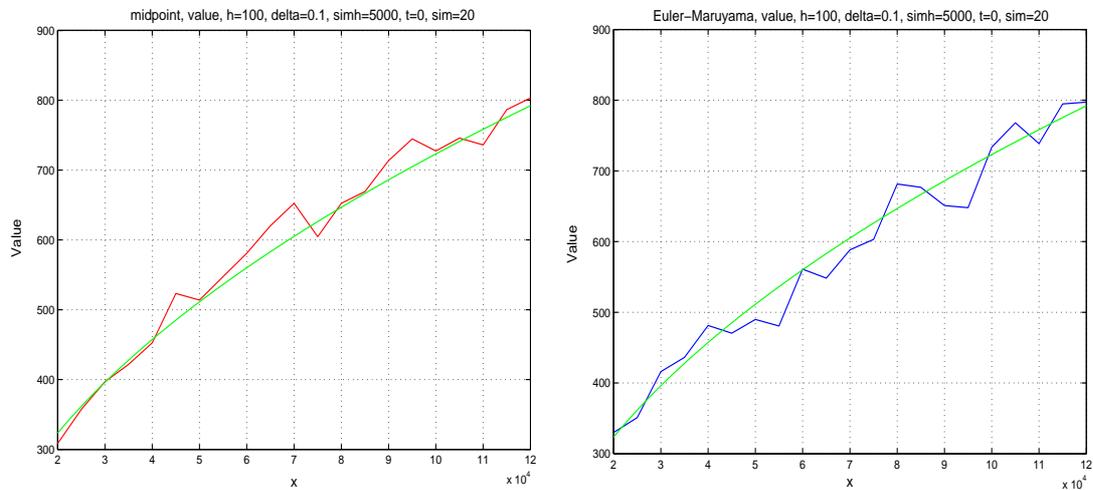


Figure 8: Expected value created by the midpoint rule (left) and the E-M method (right), in 20 simulations.

To avoid this and make the curves clearer for observation, we take $h1 = 5000$ and perform 1500 simulations for the expected value. The such obtained figures are the following.

Fig. 7 illustrates the expected optimal value from the midpoint rule, the $p-c^{10}$ method, and the Euler-Maruyama method, respectively. The theoretical solution is drawn with green solid line. It can be seen that the expected optimal value from the midpoint rule lies mainly above the theoretical line, whereas that from the Euler-Maruyama method below it. The *mixed* method of the two, the $p-c^{10}$ method of Euler-Maruyama and midpoint rule, produces then a somehow 'balanced' line along the theoretical solution, especially as the initial ($t = 0$) wealth grows large.

To observe the expected optimal value with respect to more situations of initial wealth amount, we enlarge the interval of x to $[20000, 120000]$ in the next experiment.

In Fig. 8, 20 simulations are performed for the expected optimal values. Other data

are the same with that for Fig. 7. In Fig. 8, the left is for the midpoint rule, and the right for the Euler-Maruyama method. No obvious violation of the tendency discovered on [80000, 120000] of the behavior of the numerical methods is observed in this larger interval [20000, 120000].

6. Conclusion

Theoretical and empirical analysis show that, the midpoint rule and p - c^k discretisation of the state equations for Markov chain approximation are applicable and may provide with better simulation of feedback control rules for stochastic control problems than the Euler-Maruyama method. The expected optimal values resulted from the two methods are comparable, for the aforementioned portfolio model, with that from the Euler-Maruyama discretisation which is shown to be effective by various applications.

It would be necessary to test the behavior of these two methods through more practical models. Theoretical analysis on numerical stability of these methods would be needed. On the other hand, it might be possible to adapt other numerical techniques for solving SDEs to the state equations methods, among which the two discussed serve only as examples. Nevertheless, it is shown that, proper numerical treatment of the state equations of stochastic control problems can improve accuracy of the approximating Markov chains.

Acknowledgments The first author is supported by the China Postdoctoral Science Foundation (No. 20080430402). The authors acknowledge the kind supply of the MATLAB package SOCSOL4I for solving stochastic optimal control problems by Professor Dr. J.B. Krawczyk of the Victoria University of Wellington, New Zealand, as well as the valuable suggestions from him and the referees.

References

- [1] J. AZZATO AND J.B. KRAWCZYK, *SOCSol4I an improved MATLAB package for approximating the solution to a continuous-time stochastic optimal control problem*, Munic Personal RePEc Archive, Paper No. 1179, 2006.
- [2] J. AZZATO AND J.B. KRAWCZYK, *Applying a finite-horizon numerical optimization method to a period optimal control problem*, *Automatica*, 44 (2008), pp. 1642-1651.
- [3] W.H. FLEMING AND R.W. RISHEL, *Deterministic & Stochastic Optimal Control*, Springer-Verlag, New York etc., 1975.
- [4] L.M. HOCKING, *Optimal Control-An Introduction to the Theory with Applications*, Clarendon Press, Oxford, 1991.
- [5] P.E. KLÖDEN, *A brief overview of numerical methods for stochastic differential equations*, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.8.7565>
- [6] P.E. KLÖDEN AND E. PLATEN, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin Heidelberg, 1992.
- [7] J.B. KRAWCZYK, *A Markovian approximated solution to a portfolio management problem*, <http://www.item.woiz.polsl.pl/issue/journal1.htm>, *Inf. Technol. Econ. Manag.*, 1 (2001).

- [8] J.B. KRAWCZYK, *On loss-avoiding payoff distribution in a dynamic portfolio management problem*, The Journal of Risk Finance, 9 (2008), No. 2, pp. 151-172.
- [9] H.J. KUSHNER AND P. DUPUIS, *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer-Verlag, New York etc., 2001.
- [10] R.C. MERTON, *Optimum consumption and portfolio rules in a continuous-time model*, J. of Economic Theory, 3 (1971), pp. 373-413.
- [11] G.N. MILSTEIN, *Numerical Integration of Stochastic Differential Equations*, Kluwer Academic Publishers, Dordrecht, 1995.
- [12] G.N. MILSTEIN, YU.M. REPIN AND M.V. TRETYAKOV, *Numerical methods for stochastic systems preserving symplectic structure*, SIAM J. Numer. Anal., 40 (2002), No. 4, pp. 1583-1604.
- [13] C. MUNK, *Numerical methods for continuous-time, continuous-state stochastic control problems*, 1997, <http://www.sam.sdu.dk/~cmu/papers/mchain.pdf>